

Unclassified

NEA/CSNI/R(98)1



Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

OLIS : 18-Feb-1998
Dist. : 06-Mar-1998

PARIS

English text only

NUCLEAR ENERGY AGENCY
COMMITTEE ON THE SAFETY OF NUCLEAR INSTALLATIONS

NEA/CSNI/R(98)1
Unclassified

Cancels & replaces the same document:
distributed 14-Jan-1998

**CRITICAL OPERATOR ACTIONS: HUMAN RELIABILITY MODELING AND
DATA ISSUES. Principal Working Group No. 5 - Task 94-1**

**Final Task Report prepared by a Group of Experts of
the NEA Committee on the Safety of Nuclear Installations**

62060

Document complet disponible sur OLIS dans son format d'origine
Complete document available on OLIS in its original format

English text only

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Pursuant to Article I of the Convention signed in Paris on 14th December 1960, and which came into force on 30th September 1961, the Organisation for Economic Co-operation and Development (OECD) shall promote policies designed:

- to achieve the highest sustainable economic growth and employment and a rising standard of living in Member countries, while maintaining financial stability, and thus to contribute to the development of the world economy;
- to contribute to sound economic expansion in Member as well as non-member countries in the process of economic development; and
- to contribute to the expansion of world trade on a multilateral, non-discriminatory basis in accordance with international obligations.

The original Member countries of the OECD are Austria, Belgium, Canada, Denmark, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The following countries became Members subsequently through accession at the dates indicated hereafter: Japan (28th April 1964), Finland (28th January 1969), Australia (7th June 1971), New Zealand (29th May 1973), Mexico (18th May 1994), the Czech Republic (21st December 1995), Hungary (7th May 1996), Poland (22nd November 1996) and the Republic of Korea (12th December 1996). The Commission of the European Communities takes part in the work of the OECD (Article 13 of the OECD Convention).

NUCLEAR ENERGY AGENCY

The OECD Nuclear Energy Agency (NEA) was established on 1st February 1958 under the name of the OEEC European Nuclear Energy Agency. It received its present designation on 20th April 1972, when Japan became its first non-European full Member. NEA membership today consists of all OECD Member countries except New Zealand and Poland. The Commission of the European Communities takes part in the work of the Agency.

The primary objective of the NEA is to promote co-operation among the governments of its participating countries in furthering the development of nuclear power as a safe, environmentally acceptable and economic energy source.

This is achieved by:

- *encouraging harmonization of national regulatory policies and practices, with particular reference to the safety of nuclear installations, protection of man against ionising radiation and preservation of the environment, radioactive waste management, and nuclear third party liability and insurance;*
- *assessing the contribution of nuclear power to the overall energy supply by keeping under review the technical and economic aspects of nuclear power growth and forecasting demand and supply for the different phases of the nuclear fuel cycle;*
- *developing exchanges of scientific and technical information particularly through participation in common services;*
- *setting up international research and development programmes and joint undertakings.*

In these and related tasks, the NEA works in close collaboration with the International Atomic Energy Agency in Vienna, with which it has concluded a Co-operation Agreement, as well as with other international organisations in the nuclear field.

© OECD 1998

Permission to reproduce a portion of this work for non-commercial purposes or classroom use should be obtained through Centre français d'exploitation du droit de copie (CCF), 20, rue des Grands-Augustins, 75006 Paris, France, for every country except the United States. In the United States permission should be obtained through the Copyright Clearance Center, Inc. (CCC). All other applications for permission to reproduce or translate all or part of this book should be made to OECD Publications, 2, rue André-Pascal, 75775 PARIS CEDEX 16, France.

COMMITTEE ON THE SAFETY OF NUCLEAR INSTALLATIONS

The Committee on the Safety of Nuclear Installations (CSNI) of the OECD Nuclear Energy Agency (NEA) is an international committee made up of senior scientists and engineers. It was set up in 1973 to develop, and co-ordinate the activities of the Nuclear Energy Agency concerning the technical aspects of the design, construction and operation of nuclear installations insofar as they affect the safety of such installations. The Committee's purpose is to foster international co-operation in nuclear safety among the OECD Member countries.

The CSNI constitutes a forum for the exchange of technical information and for collaboration between organisations which can contribute, from their respective backgrounds in research, development, engineering or regulation, to these activities and to the definition of the programme of work. It also reviews the state of knowledge on selected topics on nuclear safety technology and safety assessment, including operating experience. It initiates and conducts programmes identified by these reviews and assessments in order to overcome discrepancies, develop improvements and reach international consensus on technical issues of common interest. It promotes the co-ordination of work in different Member countries including the establishment of co-operative research projects and assists in the feedback of the results to participating organisations. Full use is also made of traditional methods of co-operation, such as information exchanges, establishment of working groups, and organisation of conferences and specialist meetings.

The greater part of the CSNI's current programme is concerned with the technology of water reactors. The principal areas covered are operating experience and the human factor, reactor coolant system behaviour, various aspects of reactor component integrity, the phenomenology of radioactive releases in reactor accidents and their confinement, containment performance, risk assessment, and severe accidents. The Committee also studies the safety of the nuclear fuel cycle, conducts periodic surveys of the reactor safety research programmes and operates an international mechanism for exchanging reports on safety related nuclear power plant accidents.

In implementing its programme, the CSNI establishes co-operative mechanisms with NEA's Committee on Nuclear Regulatory Activities (CNRA), responsible for the activities of the Agency concerning the regulation, licensing and inspection of nuclear installations with regard to safety. It also cooperates with NEA's Committee on Radiation Protection and Public Health and NEA's Radioactive Waste Management Committee on matters of common interest.

* * * * *

The opinions expressed and the arguments employed in this document are the responsibility of the authors and do not necessarily represent those of the OECD.

Requests for additional copies of this report should be addressed to:

Nuclear Safety Division
OECD Nuclear Energy Agency
Le Seine St-Germain
12 blvd. des Iles
92130 Issy-les-Moulineaux
France

FOREWORD

The NEA Committee on the Safety of Nuclear Installations believes that an essential factor to achieving their mandate is the continuing exchange and analysis of technical information. To facilitate this exchange the Committee has established working groups. Principal Working Group No. 5 (PWG5) on Risk Assessment was established in 1982. The mandate of this group states: "The group should deal with the technology and methods for identifying contributors to risk and assessing their importance, and appropriate exchanges of information on current research." Along with this mandate the group has also endeavoured to develop a common understanding of the different approaches taken in risk assessment.

Traditionally, the focus of the activities of PWG5 has been on the risks related to hardware. As a recognition of the importance of human interactions and of the difficulties encountered in their treatment, Task 94-1 on "Critical Operator Actions: Human Reliability Modelling and Data Issues" was initiated in 1994. The work carried out by the Task Force covers:

- A summary of research activities related to Human Reliability Analysis (HRA), presently conducted in Member countries
- A review of currently used HRA techniques and their limitations
- A PSA-based survey of HRA practices and results
- An outline of emerging methods and prospective outlook for HRA

Apart from HRA modelling, particular attention has been given to data issues, to the identification of factors driving the numerical results of HRAs, and to HRA-based design and procedures modifications. The present task report includes information which will benefit utilities, regulators, researchers and practitioners who have interest in applications of state-of-the-art HRA-techniques and in current development trends.

The Task Force met several times during the execution of the task. S. Hirschberg, Paul Scherrer Institute (Switzerland) served as task leader and co-edited the task report together with V. N. Dang, Paul Scherrer Institute (Switzerland). The work represents the collective effort of the Task Force. Still, the editors would like to express particular appreciation for the extensive contributions of B. Reer, O. Sträter and J. Mertens (Germany) who drafted Chapter 5 and Appendix A of the task report, and G. W. Parry (USA) who provided summary papers on errors of commission, dependencies and the status of the ATHEANA project.

The Task Force members contributing to the report were:

P. Wilmart (Belgium)	M. Hirano (Japan)
A. Grant (Canada)	Y. Kani (Japan)
V. M. Raina (Canada)	K. Muramatsu (Japan)
M. Patrik (Czech Republic)	M. F. Versteeg (Netherlands)
P. C. Cacciabue (CEC JRC-Ispra)	T. W. Kim (Korea, R. O. K.)
G. Cojazzi (CEC JRC-Ispra)	J. Calvo (Spain)
L. Reiman (Finland)	B. Gil (Spain)
R. Virolainen (Finland)	V. N. Dang (Switzerland)
J. - M. Lanore (France)	S. Hirschberg (Switzerland)
S. Poidevin (France)	P. Meyer (Switzerland)
P. M. Herttrich (Germany)	U. Schmocker (Switzerland)
J. Mertens (Germany)	R. Andrews (UK)
B. Reer (Germany)	B. Coxson (UK)
O. Sträter (Germany)	C. H. Shepherd (UK)
A. Bareith (Hungary)	J. A. Murphy (USA)
E. Holló (Hungary)	G. W. Parry (USA)
E. Traini (Italy)	A. Ramey-Smith (USA)
M. Fukuda (Japan)	N. O. Siu (USA)

TABLE OF CONTENTS

FOREWORD	4
1. INTRODUCTION.....	11
2. TASK OBJECTIVES AND SCOPE.....	12
2.1 Objective of Task	12
2.2 Participating Countries.....	12
2.3 Types of Reactors	13
3. RELEVANT RESEARCH AND DEVELOPMENT WORK BY COUNTRY	14
3.1 Belgium.....	14
3.2 Canada	16
3.3. Czech Republic	17
3.4 Finland	17
3.5 France	18
3.6 Germany	19
3.7 Hungary	20
3.8 Italy	21
3.9 Japan	22
3.10 Korea (R.O.K.).....	24
3.11 The Netherlands	25
3.12 Switzerland	26
3.13 United Kingdom	27
3.14 United States	28
3.15 European Commission	29
4. DATA NEEDS AND SOURCES FOR HRA	30
4.1 PSA-oriented Classification of Human Interactions	30
4.2 Data Needs in View of Modeling Experiences	30
4.3 Sources of Data for Estimating Probabilities	31
4.4 Chapter References	34
5. CURRENTLY USED APPROACHES AND THEIR LIMITATIONS.....	35
5.1 Introduction to human reliability analysis (HRA) in technical systems	35
5.2 Classification of human actions and errors	36
5.3 Basic procedure in human reliability analysis	47
5.4 Identification of safety-relevant errors.....	53
5.5 Methods for quantitative human reliability analysis (HRA).....	58
5.6 The standard method: THERP	63
5.7 New developments with emphasis on time-reliability correlations in accident diagnosis	72
5.9 Requirements for a method for human reliability assessment	87
5.10 Conclusions.....	92
5.11 Chapter References	92

6. RESULTS OF HRA SURVEY	104
6.1 The Survey.....	104
6.2 Overview of Survey Responses	105
6.3 Critical Operator Actions.....	108
6.4 Detailed Treatments.....	110
6.5 Comments on “Own” Methods.....	123
6.6 HRA-based Improvements of Design and Procedures.....	135
6.7 Conclusions from the Survey.....	139
6.8 Chapter References	141
7. SPECIAL TOPICS IN HRA	142
7.1 Modeling Errors of Commission	142
7.2 Modeling Dependencies.....	153
7.3 Modeling Recoveries	156
7.4 HRA in the Context of External Events Analysis.....	158
7.5 HRA in PSAs for Low Power and Shutdown Conditions.....	160
7.6 Consideration of Organisation and Management Factors.....	161
7.7 Transferability of Simulator-Based Data	164
7.8 Chapter References	167
8. CURRENT DEVELOPMENT TENDENCIES	170
8.1 Introduction.....	170
8.2 Summary of Issues and Research Trends	170
8.3 Recently Developed and Emerging Methods.....	179
8.4 Prospective Outlook For HRA.....	184
8.5 Chapter references	188
9. CONCLUSIONS AND RECOMMENDATIONS.....	191
APPENDIX A. DESCRIPTION OF HRA METHODS.....	195
A.2 Brief Description of EDFs PHRA	203
A.3 Brief Description of HCR	207
A.4 Brief Description of HCR/ORE	210
A.5 SLIM.....	215
A.6 HEART	216
A.7 INTENT	216
APPENDIX B. TASK 94-1 SURVEY QUESTIONNAIRE.....	218
APPENDIX C. QUESTIONNAIRE ON DETAILED TREATMENTS.....	221
APPENDIX D. IMPORTANT ACTIONS (TABLES).....	223
D.1 Important Actions for BWRs (by PSA)	223
D.2 Important Actions for PWRs (by PSA).....	227
APPENDIX E. THE ATHEANA METHODOLOGY: EXTENDED SUMMARY.....	237
I. Introduction	237
II. Overview of the ATHEANA method.....	237
III. The ATHEANA process	239
IV. Summary.....	245
V. References.....	245
APPENDIX F. CONTAINED IN A SEPARATE VOLUME	

TABLES and FIGURES

Table 2-1.	Countries contributing to Task 94-1.....	13
Table 3-1.	Status of Belgian PSAs.....	14
Figure 4-1.	Human Reliability Assessment data.....	32
Figure 5-1.	Interdisciplinary procedure in predictive human reliability analysis.....	35
Table 5-1.	Human activities prior to and after commissioning a plant	36
Table 5-2.	PRA-relevant results of human activities.....	36
Table 5-3.	Errors covered by human reliability analysis (HRA), broken down according to plant state and degradation of plant safety	38
Table 5-4.	Examples of latent errors caused by the operating personnel	38
Table 5-5.	Examples of accident-initiating (active) errors of the operating personnel	38
Figure 5-2.	Percentages of different types of human errors for 51 events in nuclear power plants [Ghertman 85].....	39
Table 5-6.	Examples of unrequired operator actions from operational experience with nuclear power plants.....	40
Table 5-7.	Important operator actions included in the DRS-B accident sequence diagrams (event trees) and quantified (after [DRS-B], Section 5.2).....	40
Table 5-8.	Categories of incorrect human outputs. “Any of these incorrect human outputs may be the result of other human errors: an error of interpretation of a pattern of signals, a misreading of a display, a misprint in an emergency operating procedure, etc. In an HRA, the incorrect human outputs and human errors leading to these incorrect outputs must be analysed “ ([Swain 83], page 2-16).....	43
Table 5-9.	Error types included in human reliability analysis.....	44
Table 5-10.	Coherence of cognitive level and situational factors and involved error types	46
Table 5-11.	Main headings of categories for information deficiencies as causes for human errors. Translated from [Hacker 86] (page 435). The categories C.1, C.2 and C.4 are specified by eight sub-categories, see [Hacker 86], Section 10.2 for details.	47
Table 5-12.	Rough classification of steps recommended for human reliability analysis in the PRA Guide [PRA-PG]. The steps are to apply to each “human-related event” identified by the fault tree analyst.	48
Figure 5-3.	A framework to collect human reliability data for various HRA methods.....	52
Table 5-13.	Nine questions on data for HRA.....	53
Figure 5-4.	Deductive procedure for the identification of operator errors contributing to system failures	53
Figure 5-5.	Survey of factors influencing individual behaviour (translated from [Grote 93], page 6)	55
Figure 5-6.	Simplified model for the incorporation of human components into a man-machine system From [Swain 83] (page 4-10), slightly simplified.....	56
Figure 5-7.	Simulation of human intention formation using CES [Woods 88]	57
Table 5-14.	Methods for human reliability quantification and excerpts from their evaluations.....	59
Table 5-15.	Significant methods for human reliability quantification.	61
Table 5-16.	Classification of the considered methods into decompositional and holistic methods.....	61
Table 5-17.	Classification of the considered methods according to their data scale.....	62
Table 5-18.	Classification of the considered methods according to their key parameters.	63
Table 5-19.	Hypothetical example concerning the dependence of the scope of diagnosis on accident severity	68

Table 5-20. Comparative evaluation of HRA methods according to percentiles of the time required for the diagnosis of an abnormal event. The HCR/ORE percentiles are based on imprecise readings from ([Moieni 94], Figure 6), the response types are defined in Table A-9.....	79
Table 5.21. Example calculation for the sensitivity of a SLIM result when changing the available time.....	81
Table 5-22. The linear dependence of the logarithmic success probability ($\log(q)$) is not equivalent to the linear dependence of the logarithmic error probability ($\log(p)$); (Based on [Kosmowski 94b]).....	81
Table 5-23. The linear dependence of the logarithmized error chance ratio is equivalent to the linear dependence of the logarithmized odds ratio.....	82
Table 5-24. Quantitative comparison of HEART and SLIM.....	85
Table 5-25. Quantitative comparison between INTENT and SLIM.....	86
Table 5-26. Summary of the catalogue of requirements.....	91
Table 6-1. Survey responses listing country and PSA study.	106
Table 6-2. PSA studies (abbreviations).	106
Table 6-3. Some characteristics of surveyed PSAs and HRAs.....	107
Table 6-4. Important actions in common (BWR).	109
Table 6-5. Important actions in common (PWR).	110
Table 6-6. Treatment of “manual depressurization” in the BWR PSAs surveyed.	114
Table 6-7. Treatment of “SLIC Actuation” in the BWR PSAs surveyed.	115
Table 6-8. Treatment of “feed and bleed” in the PWR PSAs surveyed. (p. 1 of 2)	119
Table 6-9. Treatment of “feed and bleed” in the PWR PSAs surveyed. (p. 2 of 2)	120
Table 6-10. Treatment of “Alignment for Recirculation” in the PWR PSAs surveyed.	121
Table 6-11. Treatment of “Loss of RHR” in the PWR PSAs surveyed.....	122
Table 6-12. HRA-based Improvements Summarised (listed by PSA study code).	136
Figure 8-1. Principal lines of research in HRA	173
Table 8-1. Uses of data from plant simulators and from operational events.....	175
Figure 8-2. Tree representation of decision errors and their consequences.....	178
Table 8-2. Dynamic operator-plant models for HRA-related applications	186
Table 9-1. Areas of research related to human performance listed by country	191
Figure A-1. THERP event tree modelling for the failure of the task of changing from feed to circulation mode as soon as the alarm warns of an excessively low water level in the refuelling water storage tank (RWST) during a large loss-of-coolant accident.....	196
Table A-1. Important operator errors contained in the THERP database.....	198
Table A-2. Sources of Swain's handbook error probabilities	198
Table A-3. Time-dependent parameters of diagnosis failure probability.....	200
Table A-4. Guidelines for assessing the time dependence of diagnosis failure probability.....	200
Table A-5. THERP guidelines for assessing the level of dependence between two tasks or acting persons (after [Swain 83], Chapters 10 and 18).	201
Table A-6. THERP assumptions on personnel available to cope with an accident and existing levels of dependence (after [Swain 83], Table 20-4).	202
Table A-7. Variability causes to be covered by the uncertainty factors in Swain's handbook. From [Swain 83] (pages 7-9 to 7-10)].....	203
Figure A-2. Time-dependent curves of diagnosis failure probability recommended in EdF's PHRA [EPS 900] (page 80).....	205
Table A-8. Some typical values of the probability $\text{pr}(T>t a)$ that the action is not performed within t on condition that the action is performed at all (derived from [Mosneron 90] (Figure A-2). In selecting the curve, the complexity of the situation and the availability of	

experimental results must be taken into consideration [EPS 1300] (page 112).....	205
Figure A-3. Dependence of the error probability p on the quotient of available time (t) and required average time ($T_{0.5}$) as quantified in the HRC model [Hannaman 84] for the performance of a task.....	208
Figure A-4. Expanded operator action tree [Hannaman 85] for the incorporation of the HCR model. The HCR model itself (Figure A-3) serves to estimate $P(A3)$	210
Figure A-5. Generalised event tree for modelling procedure-driven operator actions in an accident ([Moieni 94], Fig. 4).....	211
Figure A-6. Decision tree for determining the probability $p_{1,g}$ that a diagnostic logic in a procedure is misinterpreted ([Moieni 94], Figure 7).....	212
Table A-9. Standard deviation (σ) of the logarithmic diagnosis time ($\ln(t)$) as a function of response and reactor type. After [Moieni 94] (Figures 3 and 6). Except for PWR response type 3, the curves are plotted together with their uncertainty bounds.	214

1. INTRODUCTION

The treatment of human interactions is considered one of the major limitations in the context of Probabilistic Safety Assessment (PSA). While the results of many PSAs show a very significant contribution of human errors, large uncertainties are normally associated with the quantitative estimates of these contributors. This problem becomes even more significant when analysing human interactions under special conditions, for example in accident scenarios for external events or for the shutdown and low power conditions. Any improvement in the current state of knowledge with respect to the data for human interactions would have a positive impact on the confidence in PSA results, including correct ranking of the dominant accident scenarios. At the same time many PSAs have been successful at identifying critical operator actions; in most cases the benefits of these qualitative insights are not jeopardised by lack of numerical precision in the estimates.

The present HRA approaches as generally applied in PSAs are also limited in scope; for instance, they either ignore errors of commissions or treat these superficially. New, dynamic methods, primarily aiming at the resolution of the issues of cognitive errors including errors of commission are emerging but their full-scope applications within the PSA framework belong to the future.

In the context of data, some progress has been observed partially due to use of simulators to support the human reliability analysis (HRA). These applications have been rather concentrated (but not limited) to France and USA. Recently, a very promising program has been established in Hungary. The experiences from such applications are not widely known and dissemination of the relevant insights to the PSA community has some definite merits.

With respect to the identification of critical operator actions there is in some cases clear evidence and in others a good potential that the existing PSA studies may provide useful, partially generic information. The same may apply to the experiences made in the context of design and procedures improvements, based on or related to HRA.

As a recognition of the importance of human interactions and of the need to exchange experiences from their treatment, Task 94-1 was initiated within PWG5 in 1994.

The present report summarises the results of the work carried out by the group of experts. In Chapter 2 the specific task objectives are stated and the scope is defined. Chapter 3 contains the descriptions of the current HRA activities, including both industrial applications and research projects, in the countries participating in the task. In Chapter 4 data needs and sources for HRA are outlined and in Chapter 5 currently used analysis approaches and their limitations are discussed. Results of the HRA survey, carried out as a major part of this task, are presented in Chapter 6. Chapter 7 deals with a number of special topics in HRA, considered as particularly complex and/or difficult due to the scarceness of data. Current development tendencies are addressed with considerable detail in Chapter 8, followed by conclusions and recommendations (Chapter 9). Comprehensive references are provided at the end of each chapter. Finally, Appendices B, C, D, and F contain detailed information related to the HRA survey.

2. TASK OBJECTIVES AND SCOPE

2.1 Objective of Task

The focus of the task is on dynamic operator actions, which represent the most serious modeling challenge due to the scarcity of data. According to most PSAs, among the different categories of human interactions, they normally also have the highest safety significance. Interactions that occur prior to the initiating event or that lead to a plant transient are not ignored within this task but have a lower priority. Moreover, the emphasis is on the actions related to accident prevention (operator actions taken in response to initiating events); accident mitigation (accident management actions) may be included in the future but will probably be considered for a subsequent task after the completion of the present one.

The specific objectives are:

1. Based on a number of available PSAs, to identify critical operator actions during accident conditions, with particular emphasis on interactions of potentially generic interest.
2. To survey the probabilities assigned to selected interactions and the approaches used for generating these data.
3. To consider the potential for improvements of the current modeling and data situation through wider use of already established approaches as well as novel approaches.
4. To identify examples of major modifications in procedures and design resulting from PSA findings that relate to operator actions.

In order to meet these objectives the activity was divided into two parts, i.e. a survey of PSA-identified critical operator actions and a state-of-the-art review including development trends and needs.

2.2 Participating Countries

Table 2-1 lists the countries contributing to Task 94-1. In addition, the European Commission participated in the task and provided contributions through the Joint Research Centre ISPRA. The organisations involved in the work included authorities, vendors, utilities and research institutes.

Table 2-1. Countries contributing to Task 94-1.

Belgium
Canada
Czech Republic
Finland
France
Germany
Hungary
Italy
Japan
Korea (R.O.K.)
Netherlands
Spain
Switzerland
United Kingdom
United States

2.3 Types of Reactors

In addition to PSAs for Boiling Water Reactors (BWRs) and Pressurised Water Reactors (PWRs), PSAs for advanced reactor designs and heavy water reactors are also included in the survey responses. While the discussions of the methodology used for the PSAs for these reactor types are encompassed within the scope of this report, the comparative discussion of results, e.g. actions identified as important contributors, is limited to the light water reactor types.

3. RELEVANT RESEARCH AND DEVELOPMENT WORK BY COUNTRY

This chapter presents summaries of the status of HRA and of current HRA-related research prepared by each participating country (organisation).

3.1 Belgium

3.1.1 General context

In Belgium, each nuclear power unit must be re-examined after ten years of operation from the viewpoint of safety. The objective of this compulsory review is to compare the actual safety level of the unit with the safety level which would result from the application of the rules existing at the time of review.

In this context, it was deemed desirable to perform probabilistic type of safety assessment to support decision in 10-year plant backfitting process.

The following table summarises general information on the status of Belgian PSAs.

Table 3-1. Status of Belgian PSAs

Unit	Type - Vendor - No. of loops - Redundancy	Start to operate	Output Power	Current PSA Status		Level 2 scope
				Analysis	Review	
Doel 1	PWR/Twin - W - 2 - 2	1974	412 MW	Completed	Started	quant. CFM*
Doel 2	PWR/Twin - W - 2 - 2	1975	412 MW			
Doel 3	PWR - FRA - 3 - 3	1982	1020 MW	Completed	Completed	qual. CFM*
Doel 4	PWR - W - 3 - 3	1985	1056 MW	Started	-	none
Tihange 1	PWR - W - 3 - 3	1975	976 MW	Completed	Started	quant. CFM*
Tihange 2	PWR - FRA - 3 - 3	1982	970 MW	Completed	Completed	qual. CFM*
Tihange 3	PWR - W - 3 - 3	1985	1065 MW	Started	-	

*: "qual. CFM" means that only a qualitative containment failure mode analysis is carried out, i.e., accident progression modelling on dominant level 1 sequences;

"quant. CFM" means that a quantitative (probabilistic) analysis is carried out.

3.1.2 HRA Activities

In these PSAs, the human reliability analysis has been widely investigated.

Accident initiating errors are supposed to be taken into account in the probability of initiating events. Therefore, no detailed HRA has been performed to quantify this type of error.

Three types of pre-accidental activity may lead to human errors:

- component maintenance;
- periodical tests;
- locking of status changes.

Test and maintenance procedures have been analysed in order to identify the errors which contribute to reducing the availability of the system concerned by test or maintenance.

The probabilities relating to the various recovery factors are based on the values proposed by the NUREG CR-4772 : Accident Sequence Evaluation Program (ASEP) human reliability analysis procedure, Swain.

The method for post-accidental human actions is based on the THERP methodology, amended by French specific data.

Post-accidental HRA mainly consists of analysis of accidental procedures.

Errors of omission and commission and recovery actions have been performed for power states and for low power and shutdown operation modes.

HRA has led to several procedural improvements, specially in reactor shutdown states.

3.1.3 MMI Activities

For many years, Belgium Nuclear Industry has invested in high level process supervision systems aiming at helping operators deal with the large amount of information available in the control room.

Recently, a new generation of process supervision systems, called DIMOS (DIstributed MOnitoring System) has been developed by TRACTEBEL and installed in units 1, 2, 3 and 4 of the Doel plant between 1991 and 1996.

The objectives of DIMOS is to handle the raw information available within the plant in order to help the operators monitor its operation. Great focus has been put on the ergonomics aspects of the design. The Man-Machine Interface enables optimal visualisation of the right information in the right place. It is therefore distributed throughout the user's work areas (control room, technical support centre, chemistry laboratory, ...). The connection capability of DIMOS with the other computerised operator support systems available in the plant is reached thanks to an open architecture. DIMOS is also a configurable system which allows its customisation for a specific installation and its modification during the plant lifetime. The future trends in the field of operator support functionalities are directed towards alarm reduction and automatic sequence supervision.

On the other hand, developments in the area of Artificial Intelligence technologies have been carried out. Research work is in progress in the field of early warning of plant malfunctions using Artificial Neural Networks. In the framework of the HALDEN Project, a knowledge-based prototype system, tested on training simulator, is being developed aiming at helping operators deal with emergencies.

3.2 Canada

In Canada, the development of models for the quantification of operator error has evolved from the very simple models used in the early probabilistic studies of CANDU designs called the Safety Design Matrices (SDMs) to more elaborate ones developed for Ontario Hydro's risk assessments and AECL's PSA studies of research reactors.

The human reliability model used in the SDMs was a three-stepped time-based model in which the probability of a post-accident human error was taken to be 1 if available time was less than 15 minutes, 0.01 if between 15 and 30 minutes, and .001 if greater than 30 minutes. These early studies did not dwell much on pre-accident human errors, mainly on the expectation that they could be considered adequately included in equipment failure data.

A detailed human interaction taxonomy was developed during the conduct of the Darlington Probabilistic Safety Evaluation (DPSE) in the early 1980s to characterise the various types of human interactions. As well, quantification models were developed to obtain preliminary, or screening, estimates of both pre- and post- initiating event human error probabilities. The pre-initiating event models took into account the location where the error was postulated to occur (e.g., whether in the main control room, or field areas), the nature of the indications available to advise the operator of the occurrence of the error, and the likelihood of error recovery by either inspection of control room panels, field walkarounds, or periodic tests. Post-initiating event models considered the available time for action, indications received, and level of familiarity with the task. Basic HEPs were derived using the THERP methodology and data provided in the Human Reliability Handbook by Swain and Guttmann.

The DPSE methodology also comprised expert judgement elicitation and its incorporation into the final quantification of human error events determined to be the most critical by the application of the preliminary models. A paired comparison exercise was conducted involving a group of plant operators, in which they considered all possible randomly-ordered pairs of a selected number of human errors, and provided their judgement as to which event of each pair was more probable. Analysis of the paired comparisons gave a scaled relative likelihood measure, which was converted to an absolute probability scale by providing separately modelled values, based on THERP, for the events at the two extremes.

Ontario Hydro's subsequent PSAs, such as the Pickering A Risk Assessment, the Bruce B risk assessment, and the Bruce A risk assessment have generally followed the DPSE methodology. However, improvements have been made in the application of the preliminary models by simplifying the models, making them plant-specific (e.g., accounting for electromechanical indicators vs. indicating lamps), and automating the quantification and report-generation process.

AECL's Chalk River Laboratories developed in 1995 a human error probability quantification method for use in PSA studies of the NRU research reactor upgrades. Among its attributes were explicit consideration of performance shaping factors such as task unfamiliarity, design mismatch leading to misleading or ambiguous representation on control panels of plant conditions, lack of job aids such as schematics and checklists, poor feedback from actions, poor procedures, lack of checking, and information overload. It also incorporated a procedure for accounting for dependencies between human error events. The values of basic HEPs were determined based on a review of the various models described in the literature such as SHARP, HEART and THERP.

3.3. Czech Republic

In Czech Republic, the methods of probabilistic safety assessment were used to analyse the operational risk of the following subjects:

- nuclear power plant Dukovany (PWR reactor of WWER-440 type, four units, operated since 1985, PSA Level-1 opened in 1989 and finished in 1995, shutdown PSA opened in 1996)
- nuclear power plant Temelín (PWR reactor of WWER-1000 type, two units, under construction, operation planned to start in 1998, PSA Level-2 started in 1993 and finished in 1995)
- research reactor LWR-15 in Nuclear Research Institute Rez (reactor of nominal power of 10MW, operated since 1959, PSA project opened in 1995 and finished in 1996)
- chemical plant Spolana Neratovice (analysis of the most important risk contributors - production of polyvinyl-chlorid, analysis opened in 1995 and finished in 1996).

All the above mentioned projects were carried out by the specialists from Nuclear Research Institute Rez (the Temelín PSA in co-operation with NUS Halliburton and EGP Prague). In all cases, human reliability analysis was carried out in detail and human factor has been found to be one of the most important contributors to the risk of operation of the given technological unit.

Section 6.5.2 describes the methods used in the Dukovany and Temelín PSAs and discusses the experiences with the methods.

In the case of PSAs for research reactor and chemical plants, the spectrum of human interventions important from point of view of risk has been much more variable and has not been limited by procedure-driven actions performed from control room. In those cases, a broad spectrum of methods of quantification was used to address the different types of potential human failures (THERP, ASEP, HEART).

In the very next future, human reliability analyses in the frame of shutdown PSA for NPP Dukovany will be carried out using the similar methods as for the nominal power PSA (adjusted, if necessary). The possibility of using the results of simulator exercises to refine the probabilities of human failures obtained by means of above mentioned general methods in frame of NPP Dukovany PSA is being studied. Based on the current status of knowledge, this idea seems to be very promising and a new project is planned to be opened in 1997.

3.4 Finland

In 1984 the Finnish regulatory body (STUK) required the utilities to make plant-specific PSA studies for each plant. Results of Level-1 studies were submitted for regulatory review in June 1989. PSA for low power and shutdown operation modes is completed for the TVO plant and is underway for the Loviisa plant. In connection with these studies, advancements and plant specific modifications have been made to some state-of-the-art techniques in human reliability analysis (HRA).

In the PSA of the TVO plant a team work technique, where principles of HAZOP were modified, was used to identify potential maintenance and testing errors. The identified errors were quantified with

support of the Handbook data and the MSF model. In the analysis of operator errors, probabilities were assigned using the HCR model.

In the PSA of Loviisa plant, the pre-accident errors were quantified using a shortened version of the THERP procedure. The probability of repeating an error was quantified using an approach that was modified from the Handbook method.

As concerns the post-accident errors of the operators, the slow diagnosis and misdiagnosis probabilities were assigned separately in the Loviisa plant. The slow diagnosis probability was evaluated using plant-specific time reliability curves which were based on simulator experiments and expert judgement. The probability of misdiagnosis was estimated using an asymmetric confusion matrix.

In connection with the reviews of the Level-1 PSAs, STUK has made some own analyses. In these analyses SLIM, a modified HCR model and a model based on plant specific time reliability curves (TRC-SI-model) were used to analyse operator errors. A Paired Comparisons exercise was conducted to assess human errors and their dependence in test and maintenance activities. The methods used by STUK were described in the first meeting of PWG5 Task 15.

In the analysis of low power and shutdown modes for the TVO plant some other methods were used. For example, the task confusion matrix method was used in the identification and the logical modelling of hazardous combinations of different operational, testing and maintenance tasks.

At the moment STUK is starting a project the purpose of which is to analyse operating experiences of Finnish NPPs with special emphasis on maintenance errors, their dependence and possibilities to prevent them.

3.5 France

3.5.1 Utility (EDF)

1300 MWe series: up-dating of the PSA, especially taking into account the new symptom-oriented accidental procedures.

1400 MWe series (N4): development of a PSA which takes into account the specifics of the N4 series, in particular, the computerised control room and the computerised symptom-oriented procedures.

Each project includes:

- data collection, by specific simulator experiments,
- development of appropriate models,
- integration in the accident sequences.

3.5.2 Safety Authorities (CEA/IPSN)

900 MWe series: up-dating of the PSA, especially due to changes in the operating teams organisation for which new models must be developed. Moreover, a program is starting for the definition and modelling of human reliability in the fire PSA and in the level 2 PSA.

3.6 Germany

3.6.1 Activities of the GRS

3.6.1.1 HF Activities at the GRS

Up to 1992, GRS mainly performed different Precursor Studies and the German Risk Study. In 1992, the human factor research was re-established with four projects related to Human Reliability (HR) and Man-Machine Interaction (MMI).

In the HR topic, a method for the evaluation of plant experience was developed and a comparative investigation of various HRA Methods was made. The evaluation of plant experience is an ongoing activity at GRS. It concentrates on the evaluation of HR data like errors of omission and errors of commission for PSA demands. Special interest is given to identification and assessment of human cognitive errors. This task is performed in a research group together with EDF (France), the Jülich Research Center (Germany) and KEMA (Netherlands). GRS is also heavily involved in different HRA reviews and actually intends to update the German HRA guidelines.

In the MMI topic, one research project to assess the utilisation of New Information Technologies in Nuclear Power Plant (NPP) Control Rooms was recently finished. Additionally, ergonomic investigations in the test Control Room of the GRS has been performed to improve process control systems. In another investigation, the ergonomic design of Accident Management procedures is examined. Organisational aspects of crisis management were considered. Presently, methods are developed to assess the reliability of Accident Management (AM) procedures for boiling water reactors.

Currently, investigations are planned to assess human reliability in emergency situations and to assess the cognitive demands of different types of process control.

3.6.1.2 HRA Activities at the GRS

HRA of plant procedures

- Optimising existing procedures
- Developing new procedures

Computer-aided tools

- Conventional programming
- Expert system programming

Chemical plant applications

Methods for quantifying time, organisation, backfitting measures and decision-based errors

Evaluation of incident reports involving human factors

Review of methodological details

3.6.2 HRA Activities at the Research Centre Jülich (FZJ)

The aim of FSJ work is the more realistic treatment of human errors in PSA by further development and modification of THERP with respect to psychological mechanisms that are inadequately covered.

The present activities are:

- Investigations on “Unrequired” actions caused by cognitive errors in co-operation with KEMA and GRS,
- Investigation of psychological environment of operator tasks (e.g. by interviewing trainers at the NPP simulators in Essen/Germany),
- Quantification of probabilities in decision processes (handling of non-stochastic uncertainties using fuzzy set theory),
- Application of THERP to operator actions in chemical plants,
- Development of a small expert system modelling THERP (to reduce uncertainties. caused by subjectivities in using THERP).

Effort:

Approximately 3-4 man-years per year (including 2 PhD researchers).

3.7 Hungary

3.7.1 Uses of HRA

A Level-1 PSA study was completed for NPP Paks, Unit 3 in 1994 as part of the AGNES safety re-evaluation project. Subsequently, similar PSAs were performed for Units 1 and 2 in the framework of the Periodic Safety Review of the plant. All PSAs were conducted by VEIKI Institute for Electric Power Research. To reflect the most important safety features of the VVER-440/213 reactors of Paks much attention was paid to human reliability analysis in the PSA.

Pre-accident errors were analysed by using a modified ASEP procedure that took into account the specifics of VVER operation. The study included a number of talk-through analyses with extensive involvement of plant technical staff. A formalised procedure was adapted to collect information important to HRA and organise this information into a framework useful for quantitative assessment. The basic HEP estimates were derived mainly from the THERP Handbook and from the ASEP procedure guide.

Initiator type errors were quantified by the use of plant specific data. The majority of these initiators were found to be those leading to an inadvertent reactor trip.

Observations were made at the full scale replica simulator of the plant to help modelling and quantification of post-accident errors. Although the simulator study was originally aimed at developing and using plant specific TRC curves, the results of the data analysis showed that use of such curves were not sufficient. This finding led to the development of a stand-alone model of crew reliability. This model is based on a decision tree approach, and it relies on both simulator data and expert opinion. The decision tree approach was applied in the Level-1 PSA.

3.7.2 *Developments in HRA*

Currently the low power and shutdown PSA is underway for NPP Paks Units 1 and 2. This analysis also puts emphasis on human reliability assessment because the context in which human errors can occur and the conditions of human responses to plant disturbances are different from that of power operation. Extensions have been made to the approaches used for the full power PSA, which includes an HRA oriented review of test and maintenance practices, analysis of reported plant operational events, and further developments in the decision tree model for the purpose of modelling post-accident errors in low power and shutdown modes.

In addition to the low power and shutdown PSA, current HRA related efforts include methodological developments in modelling cognitive errors as well as integration of various types of data and expert opinion in support of HEP quantification. Also, precursor studies are being performed for the Hungarian regulatory body by using the plant PSA model. As far as human reliability is concerned, these studies are envisaged to result in

- a better understanding of human behaviour and
- improvements in the probabilistic description of human actions.

New emergency operating procedures are under development for Paks. The application of human reliability and human factors disciplines will probably be an important aid in introducing the new procedures at the plant.

3.8 Italy

3.8.1 *HRA Activities*

Since 1984, ENEL has performed a number of PSAs for all of the Italian Nuclear Power Plants and for the new passive plants designs being developed in the United States and in Sweden.

In all of these studies, ENEL performed the Human Reliability Analysis mainly based on United States developed methods (THERP by Swain and Guttman and the EPRI method based on simulator exercise).

For operating plants, numerous control room personnel interviews were performed in order to correctly model the mental behaviour of the control crew. Other methods have been investigated including HCR correlation for cognitive errors evaluation.

Guidelines, for conducting the Human Reliability Analyses have been developed for the new reactors analyses. The basic SHARP (Systematic Human Action Reliability procedure) framework was adopted.

An in-depth analysis on nuclear power plant operating experience in the field of human behaviour is in progress to determine root causes, direct causes, etc., of the most significant events.

3.8.2 Man-Machine Interfaces

Presently, the following activity phases are in progress or planned:

- design, construction and validation of advanced static displays for advanced (completely informatised) man-machine interfaces: in particular, systems synoptics, procedures, alarms using a software (Picasso code) developed by the Halden Reactor Project (Halden, Norway) ;
- animation of displays through simplified simulation codes;
- construction of a workstation for ergonomic; studies;
- display-workstation integration.

All of the above activities will be performed with the contribution of Nuclear Power Plants Control room operators and supervisors.

3.9 Japan

The importance of human reliability analysis is highly recognised and many new programs have been started in Japan: The programs at JAERI (Japan Atomic Energy Research Institute), NUPEC (Nuclear Power Engineering Corporation) and CRIEPI (Central Research Institute of Electric Power Industry) are comprehensive ones and cover various aspects such as human reliability analysis, man-machine interface research, operational management, training, utilisation of artificial intelligence for operational aid and collection and analysis of human reliability data. Utilities are collecting and analysing human behaviour data at their training centres. In Japan, however, HRA methods fundamentally based on the THERP (Technique for Human Error Rate Prediction) method are used in recent PSAs, because the above-mentioned programs of human reliability analysis are still on going and their results have not yet systematically been formulated for practical use in PSA.

For example, utilities used variants of the THERP method in their IPEs (PSAs on all NPPs to examine candidates for accident management) done during 1992-1994 with consideration of expert judgement on operational management in Japanese NPPs.

For human reliability analysis, JAERI developed a DEBDA methodology based on Detailed Block Diagram Analysis method. PNC has developed a PC-based interactive human reliability analysis code HURASS/SHERI principally based on the THERP method. NUPEC has conducted some level-1 PSAs on Japanese typical PWRs and BWRs using fairly exact THERP method with various expert judgements on the uncertainty of evaluation on diagnoses and actions done by operator team during accident progression.

3.9.1 *Summaries of HRA-related research activities*

Research at the Human Factors Research Laboratory of JAERI places emphasis on better understanding of characteristics of human behaviour. In order to better understand characteristics of actual operators' cognitive behaviour coping with abnormal conditions, experimental and analytical research with reactor simulator has been performed. Based on knowledge from both these empirical studies and the outcome of cognitive science/engineering, investigations made on basic characteristics of human cognitive behaviour including mechanism of human error.

Taking account this knowledge, a research program is being carried out to establish a systematic methodology for man-machine system evaluation, which includes development of a dynamic computer simulation model of man-machine system as a tool.

In order to assist in human reliability analysis for probabilistic safety assessment, a supporting tool is under development.

Institute of Human Factors (IHF) of NUPEC promotes the basic research on human factors aiming at the reduction of human error in operation and maintenance of nuclear power plants. It has the following programs:

- Human behaviour and its modelling

IHF is developing the simulation model representing human cognitive information processing for the understanding of human characteristics and the investigation of causes of human errors.

- Development of human reliability analysis methods

IHF is developing human reliability analysis methods for evaluating human response time and error rates for applying to probabilistic risk analysis and others.

- Establishment of human reliability data base

IHF is performing the collection of human reliability data and the analysis of human error events for contributing to the study on the prevention of human errors.

- Human characteristic experiments for nuclear power plants

IHF has just started since last spring various experiments relating to human behaviour and characteristics using human characteristic experimental facilities in order to verify the results of theoretical investigations carried out in IHF so far.

Research at the Human Factors Research Centre (HFC) of CRIEPI has been devoted to the Japanese utility, emphasising on developing various systems for the reduction of human errors in operation and maintenance activities at nuclear power stations.

Those systems include the Japanese version of the Human Performance Enhancement System (J-HPES) originally developed at US's INPO, the Human Errors and Behaviour Prediction System (HEBPS) and various human factors database systems.

The research activities covered in HFC/CRIEPI are: human error mechanism analysis; nuclear operator and its team behaviour modelling during cognition, decision, conversation and action, measurements on physiological and psychological work loads, maintenance work support tool development, sociological approach on human/environment relations, etc.

3.10 Korea (R.O.K.)

3.10.1 Wolsong PSA

Korea has 10 operating nuclear power units (NPPs) and 6 units under construction at 4 sites. Among them, 4 units, which are located in Wolsong site, are CANDU 600 PHWR type. One of them (Wolsong Unit 1) is operating and three Wolsong Units 2/3/4 are under construction. The internal Level 1 PSA of Wolsong Units 2/3/4 has been performed as one of conditions to achieve Operating License imposed by the Korea regulatory body, KINS (Korea Institute of Nuclear safety).

The scope and methodology being used in Wolsong 2/3/4 PSA are equivalent to those of individual plant examination (IPE) for PWR. The objective of HRA in the PSA of Wolsong Units 2/3/4 is to identify human actions related to the safety of NPPs, to quantify those human actions, and to provide information for abnormal operating procedure (AOM) writers in accordance with CANDU practice.

The EOPs of Wolsong Units 1/2/3/4 would follow present AOM of Wolsong 2 NPP. Therefore, present AOMs of Wolsong Unit 2 being prepared are reviewed for HRA in the PSA of Wolsong Units 2/3/4. The overall structure of AOMs of Wolsong Unit 2 follows the recently prepared EOPs of Pressurized Water Reactor such as YGN 3&4 in Korea.

We adopted the Systematic Human Action Reliability Procedure (SHARP) approach as an overall HRA procedure. For the detailed analysis of significant human actions, the THERP/ASEP methods are employed.

Based on the review of present AOM being prepared and the interview with Wolsong Unit 1 operators, several recommendations are given to provide information for future AOM/EOP preparation and operating staffs of Wolsong NPP. Korea has 10 operating nuclear power units (NPPs) and 6 units under construction at 4 sites. Among them, 4 units, which are located in Wolsong site, are CANDU 600 PHWR type. One of them (Wolsong Unit 1) is operating and three Wolsong Units 2/3/4 are under construction. The internal Level 1 PSA of Wolsong Units 2/3/4 is being performed as one of conditions to achieve Operating License imposed by the Korea regulatory body, KINS (Korea Institute of Nuclear safety).

3.10.2 Research Activity Related to AM HRA

Up to 1996, KAERI critically reviewed the conventional IPE/HRA methodologies (i.e. THERP, ASEP, HCR, HCR/ORE, SLIM, STAHR, etc.) and investigated the applicability of these to the human actions under accident management situations.

An approach was proposed for the evaluation of an accident management strategy when an operator action is involved. This approach classifies the failure of a selected strategy into 4 possible states: 1) the failure of an intention formation by operators, 2) the failure in taking an action following a correct diagnosis, 3) the failure of a system operation following an adequate execution, and 4) the failure due to a delayed action.

The approach considers a variability of event timing obtained from uncertainty analysis in thermal-hydraulic calculations to reflect phenomenological uncertainties. Then it uses a stochastic and sequential approach to resolve timing problems regarding operator action and system operation. The proposed method was applied to assess a cavity flooding strategy for the prevention of reactor vessel failure.

3.10.3 Simulation for Development of Group Performance Model

A simulation study for developing individual and group performance model related to realistic HRA methodology was performed at KEPCO Training Center by KINS. Full scope simulator of 950 MWe PWR type plant with SGTR scenario was utilized for empirical investigation to identify the influence factor to team performance in the point of view of relative experience and expertise level. In this investigation, eleven on-the-job main control room crew, sixty seven individuals in 950 MWe PWR type domestic plant were involved. Four influence parameters are identified to be dominant factors to team performance in main control room during the accident.

3.11 The Netherlands

For both the Dutch nuclear power plants full scope (including low power and shut-down states), Level 3 PSAs have/are performed.

3.11.1 Borssele PSA

- For the pre-accident human interactions THERP was used (e.g., miscalibration of level indicators). The dependencies between actions were taken from the Handbook (Swain & Guttman).
- The main method for HRA for the post-initiating event EOP based human actions was the so-called HCR-ORE model as developed for EPRI. This method involves models describing a cognitive/procedural error probability, a crew non-response probability and a manipulative error probability. Also manual backup to automatic actions and recovery actions were modeled in this manner. The same dependencies between actions were used as for the pre-accident human actions.
- Human actions causing an initiating event were not modeled separately but were implicitly included in the scram frequencies.
- A separate analysis of the errors of commission have been performed (power states only). This study addressed: global misdiagnosis, local errors of commission (misdiagnosis/slips); and errors contributing to initiating events. The method used was the so-called cause-effect mapping concept as developed by Dr. A. Macwan (PhD thesis Univ. of Maryland) by incorporating some recently postulated theories of cognitive mechanisms.
- An analysis of the errors of commission during the low power and shut-down states is under way.
- Human actions in the level 2 part were directly modeled in the Containment Event Trees. Some branching points reflected Accident Management Operator Actions. The error probabilities were primarily values derived via expert opinion. Actions included were e.g.: recovery of in-vessel injection (after core damage, but before or after RV failure),

depressurisation of primary coolant system (manual bleed), manual recovery of respectively electric power, containment spray and containment heat removal.

3.11.2 *Dodewaard PSA*

- For the pre-initiators, a simplified THERP method was used.
- For the post initiators, slips were modeled with THERP. Recovery errors/response errors were modeled according to SAIC (Science Applications International Corp.), time-reliability correlations (TRC). Dependencies were modeled according THERP guidelines or for multiple actions according to the SAIC AOT/TRC approach.

3.11.3 *Research*

The Dutch regulatory body, jointly with the Ministry of the Environment, is sponsoring a research project at the Delft University of Technology with a general project definition of modelling the man-machine interface. This project resulted recently in a large collaboration between Delft Technical University, CEC Joint Research Centre of Ispra, CSN (Spain), the University of Maryland and the Université Libre de Brussels. The entire collaboration is situated around DYLAM (a dynamic accident sequence simulating method) as developed by JRC-Ispra. The human operator, as will be modeled in an operator model, plays a central role in this approach.

3.12 *Switzerland*

3.12.1 *Current Status of HRA*

For the four nuclear power plants in Switzerland, three have PSAs that have been finalised and the fourth PSA is in progress. Of the three completed PSAs, two have been extensively reviewed for the Swiss Federal Nuclear Safety Inspectorate (HSK). The completed PSAs use PLG's modified SLIM methodology for quantifying dynamic (post-initiating event) human interactions. The fourth study uses a modified Accident Sequence Evaluation Programme (ASEP) methodology.

3.12.2 *Current Developments in HRA*

The Paul Scherrer Institute (PSI) is carrying out a research program to address HRA and human performance issues as they pertain to the operation and safety of Swiss nuclear power plants.

The short to medium term activities of the research program address:

- the calibration issue in the HRAs where (PLG) SLIM is applied
- the characterisation of recently proposed methods for screening studies of errors of commission (in particular, the approach applied by Parry et al.)
- the evaluation of these methods

For the longer term, the program concentrates on dynamic (simulation) approaches for addressing errors of commission and cognitive errors. In this area, one activity is the assessment of the feasibility of dynamic operator-plant models.

The PSI is interested in co-operating in efforts to resolve shared issues internationally. We are exploring issues of mutual interest and possible collaborations with other research groups.

In connection with the development of the research program, the treatments of human interactions in the Swiss PSA studies are being reviewed. The review considers the PSA results related to the human performance component of safety, HRA methodology, and comparisons among Swiss PSAs.

In addition to participating in this task of the OECD NEA Principal Working Group #5, the PSI is working in the frame of the IAEA Co-ordinated Research Programme on Collection and Classification of Human Reliability Data for Use in Probabilistic Safety Assessments. The objective of the Swiss work is to identify scenarios and operator interactions to be emphasised in data collection efforts on the basis of PSA results.

In human factors research, the HSK participates in the international man-machine interaction programme co-ordinated by the Halden Reactor Project in Norway.

3.13 United Kingdom

The UK has three main reactor types in operation, Magnox, AGR and Sizewell B (a PWR).

The Magnox plants have undergone a periodic safety review after 20 years of operation and several are already part way through a Continued Operation Review, which involves the production and assessment of a detailed plant specific PSA. The AGR plants are also undergoing Periodic Safety Reviews, involving similar PSAs, to underwrite the next ten years of operation for each station.

These studies have confirmed the reliance placed upon manual actions such as level control, pump start, etc.

Plant PSAs typically model 70-100 operator actions (Type C) using the HEART technique. Error rates vary from approximately $1E-5$ (failure to trip manually) to 0.2 (failure to isolate CW pump without procedures).

Operator errors and CCFs tend to dominate the overall risk calculations.

Further risk reduction claims are justified by reference to System Based Emergency Response Guidelines (SBERGs); component repairs or other corrective actions. Transient calculations have shown that many human recovery actions have considerable time scales within which they can be successfully achieved (i.e. >24 hrs.). Consideration of dependency between the numerous human actions modelled seen as essential.

Overall, the human factors assessments have led to a significant number of procedural improvements.

The Sizewell B PWR base case PSA is now complete. Operator actions are modelled within both the Fault and Event Trees. About 60 types of operator action (such as initiation of boration following secondary depressurisation) are modelled. Each action type is further subdivided (e.g. fail to letdown; fail to borate) such that around 100 actions are modelled with HEPs ranging from $1E-6$ to 0.2.

Operator actions are quantified using HEART, with a few exceptions where THERP seems more appropriate. Dependency is modelled using a power law. Additionally, some judgmental HEPs are applied directly to certain plant damage state sequences.

The core damage sequences in the Sizewell B base case have now been developed for running on Risk Spectrum software to support a station Living PSA.

The United Kingdom has sponsored some interesting work via Birmingham University. The study provides code to code and code to data comparisons using THERP, JHEDI and HEART for a variety of nuclear and non-nuclear tasks. The study indicates a very good level of consistency.

3.14 United States

3.14.1 Current Uses of HRA

In the United States, the industry making the most use of human reliability analysis (HRA) methods remains the nuclear industry. This largely reflects the predominant use in this industry of probabilistic risk assessment (PRA) as one factor in making safety assessments by the regulating authority, the U.S. Nuclear Regulatory Commission (NRC). Nearly all operating units in the commercial nuclear power industry have submitted PRA studies for review under the NRC's Individual Plant Examination (IPE) program. Within the IPE program guidelines, a utility was permitted to select a HRA method that it felt was appropriate. As a result, several HRA methods were used in IPE PRAs. These include the time-based Human Cognitive Reliability (HCR) method, the expert-judgement-based Success Likelihood Index Method (SLIM), the Technique for Human Error Rate Prediction (THERP), and variants of the THERP method such as the Accident Sequence Evaluation Program (ASEP) method and that developed by the Westinghouse Corporation. The draft NRC report entitled „Individual Plant Examination Program: Perspectives on Reactor Safety and Plant Performance”, (NUREG-1560, Volume 1, Part 1, Chapter 5, October 1996) summarizes in some detail the human actions important in the IPEs and addresses the degree of variability in the results of the HRAs across the different IPEs.

In addition to the commercial nuclear power industry, the facilities operated on behalf of the U.S. Department of Energy (DOE) use HRA and PRA studies for evaluating safety concerns in operations and for environmental remediation programs. HRA methods tend to be selected for the specific application and site from among standard methods, though in some cases innovative use is made of site-specific data combined with expert judgement. Other agencies of the U.S. government and private industries (particularly the chemical process and petrochemical industries) are increasing their use of HRA, though these efforts are limited and rarely published.

3.14.2 Current Developments in HRA

In the United States, essentially all research and development of HRA methods are being sponsored by the NRC. One research program is to develop and demonstrate an improved HRA method that creates a more accurate representation of human behavior in PRA studies, particularly one that accounts for the causes and consequences of errors of commission and dependencies among multiple human errors. To support this effort, detailed information from events involving human errors is being gathered and analyzed in terms of a framework of human behavior developed from psychological principles of classes and causes of human error. Another program is assessing the influences of management and organizational practices

on personnel and plant performance. Beyond these nuclear power-related programs, no substantial development work is known to be underway in HRA.

3.15 European Commission

3.15.1 Activities of the Joint Research Centre, Institute for Systems Engineering and Informatics

An extensive activity on Human Factors is carried out at the Institute for Systems Engineering and Informatics of the Joint Research Centre of the European Commission. The research programmes which support the Human Factors studies are: the Working Environment and the Nuclear Reactor Safety. The objectives of the research are to develop means to support the work of humans with complex computerised technology and to prevent accidents related to human causes.

These objective are tackled from three different perspectives:

1. The development of analytical simulations, for studying the behaviour of humans and plants during accidents. This work aims at improving the design of tools for the control of systems and the early detection of accidental events.
2. The performance of field analysis of the workplace, for studying actual working environments from a cognitive ergonomics viewpoint. This research aims at providing means for evaluating organisational factors and for promoting improvements in the design of future control rooms.
3. The development of probabilistic methods for human errors evaluation, for performing the safety assessment of plants showing substantial dynamic human-machine interaction.

These activities are carried out with particular attention given to the area of aviation and nuclear power plant safety and control, where a large portion is attributed to computerised technology, while ensuring continuous adaptation of human skill to supervising and managing the automatic systems.

In particular, the following tools have been developed:

- A cognitive model for the simulation of operators behaviour (COSIMO), for the study of accidents in the nuclear and aviation domains.
- A dynamic human reliability methodology (HERMES), for the prospective and retrospective studies of the human errors to be included in Probabilistic Safety Studies.
- An empirically based model of operators' behaviour under stress.
- A data bank (HEAT) of accidents related to human factors collected in different European industrial environments, focused on the evaluation of the socio-technical aspects of the working environment.

4. DATA NEEDS AND SOURCES FOR HRA

This chapter discusses the data needs and sources. The PSA perspective is presented in terms of a classification of human interactions (which defines the type of data needed) and in terms of relevant PSA insights (which indicates what type of data is particularly desirable in view of the relative importance for the PSA results and/or in view of the current availability and quality of specific types of data). Data requirements associated with evolutionary and revolutionary models are touched upon in Chapter 8 in the presentation of emerging methods and of the prospective outlook for HRA.

4.1 PSA-oriented Classification of Human Interactions

For PSA, a classification of human interactions based on the time of occurrence is frequently used. The PSA addresses three time frames, which are defined relative to the initiation of an abnormal scenario that may potentially lead to plant damage or an accident. These are: 1) normal operation, which includes maintenance and testing; 2) the initiation of abnormal scenarios; and 3) the response of the plant and human operators in the abnormal scenarios. Corresponding to these time frames, the categories are:

Category A: Interactions that occur prior to an event that initiates a scenario, an initiating event. Errors associated with these interactions result in equipment in standby systems being unavailable to perform their function as required should an initiating event occur.

Category B: Human interactions that initiate a scenario, also referred to as human-induced initiators.

Category C: Actions taken by plant staff after the initiating event. The staff performs these actions in response to the scenario to bring the plant to a safe state. These actions, which are also called “dynamic operator actions”, may be guided by procedures or they may not.

For incorporation into the PSA/QRA, category C interactions may be further separated into three different types:

Category C1: Procedural safety actions. These actions involve success or failure in following procedures or rules known to operators in response to an accident sequence. In the failure space these actions are referred to as errors of omission.

Category C2: Aggravating actions (also referred to as errors of commission).

Category C3: Improvising recovery/repair actions. They may include the recovery of previously unavailable equipment or the use of non-standard procedures to ameliorate the accident conditions.

4.2 Data Needs in View of Modeling Experiences

Below follow some comments on the experiences from modeling of different categories of human interactions in nuclear PSAs (Hirschberg et al., 1991). For other types of large scale industrial activities it is difficult to make any generalisations. This is due to the variety of processes and a very limited number of published applications.

Category A interactions have usually not been found among the dominant risk contributors. In applicable cases the reasons for this are the positive impact of redundancy, diversity and separation of standby safety systems. The modeling is rather straightforward, although difficulties arise when accounting for dependent failures that might lead to unavailability of component groups. Because many category A interactions, such as maintenance and testing activities, are regularly performed, plant experience can provide information. Quantitative data can be collected on human errors such as miscalibration, failures to restore valves to the correct positions after testing and maintenance, and failures to reconnect electric power after maintenance. Although obtaining the estimates for use in the PSA remains to some extent judgmental, the estimates are considered relatively reliable. Account for the impact of recovering factors requires relatively deep consideration of plant-specific aspects (procedures, layouts, status control) and is subject to larger uncertainties. However, existence of recovery factors often significantly reduces the absolute values, and consequently also the concerns about uncertainties.

Category B interactions are seldom explicitly identified in PSAs. The frequencies of the initiating events, for which relatively extensive data are available from operating experience, implicitly reflect these interactions. If a root cause analysis is performed within a PSA to identify specific causes for initiating events, it is usually rather superficial and is aimed primarily at assessing the probabilities of recovery from initiating events (those caused by human errors can sometimes be quickly recovered).

Category C interactions represent the greatest challenge. Generally data on this category of interactions are rare (because abnormal scenarios fortunately occur rarely). Most attention has been given to procedure-driven interactions belonging to Type C1, which also frequently appear in the accident sequences that dominate the risk. Type C2 interactions are most difficult to deal with and are usually included only to a very limited extent, i.e. those errors of commission that do not make the plant condition worse can be treated as Type C1 errors of omission (an example being turning the wrong switch, but with no impact on operating systems). Their probabilities have to be added to those of the corresponding errors of omission. The second class of errors of commission occurs when the operator correctly diagnoses the event, but chooses a non-optimal strategy for dealing with it. The third class of errors of commission may occur when the operator's mental image of the plant differs from the actual state leading him to perform the right actions for the wrong event. Identification of class two and three errors of commission is challenging and no proven systematic methodology has been developed, although some recent PSAs do include them on a case-by-case basis. Examples of errors of commission can be found among Licensee Event Reports, but few data exist for a full-scope analysis.

The non-procedurally driven actions are mainly Type C3 recoveries. For these actions relatively extensive data may in some cases be available, for example times for recovery of offsite power or component repair times. Data assigned to manual opening of a failed valve are judgmental and are usually not very important for PSA results.

4.3 Sources of Data for Estimating Probabilities

Data is here defined broadly to include sources of both qualitative and quantitative information. For HEP estimation, it is not only necessary to know the number of errors that have occurred for each task and how often the task was performed but also the circumstances under which the task was performed.

The circumstances for each task and the error occurrences should be compared with the PSFs for the task to be quantified. Estimates should place more weight for the data from similar circumstances. In this context, plant site visits including, whenever applicable, interviews of operators, walk-throughs of the plant, reviews of control room layout, examination of access to equipment required for recovery actions, etc., help to organise the appropriate human reliability data base.

Particularly for non-dynamic actions (e.g. maintenance actions that are quantified in the Cat. A analysis), the HRA Handbook (Swain and Guttman, 1983) still constitutes the most comprehensive source of data. The data were collected from diverse sources within and outside the nuclear industry and adjusted for typical US nuclear power plant conditions. It documents how the basic HEPs can be adjusted for differences in the PSFs and offers a practical guide to its application into the PSA framework.

For Category C, “dynamic operator” actions in particular, the availability of adequate data is very limited. As a result, multiple sources of data should be considered (e.g. Fig 4-1):

1. Data are obtained from the recording of real emergency operations as well as during training sessions, in particular at training simulators.
2. Data can be collected in the working environment and within the organisation by observation of normal operations, e.g. during everyday work in the control room, or in maintenance.
3. Data can be obtained by interviews and questionnaires at all levels of the organisation.
4. Data are contained databases derived from mandatory and voluntary reports on accidents, incidents and near misses. These include event reports, near-miss and precursor reports, maintenance reports, and plant log books.

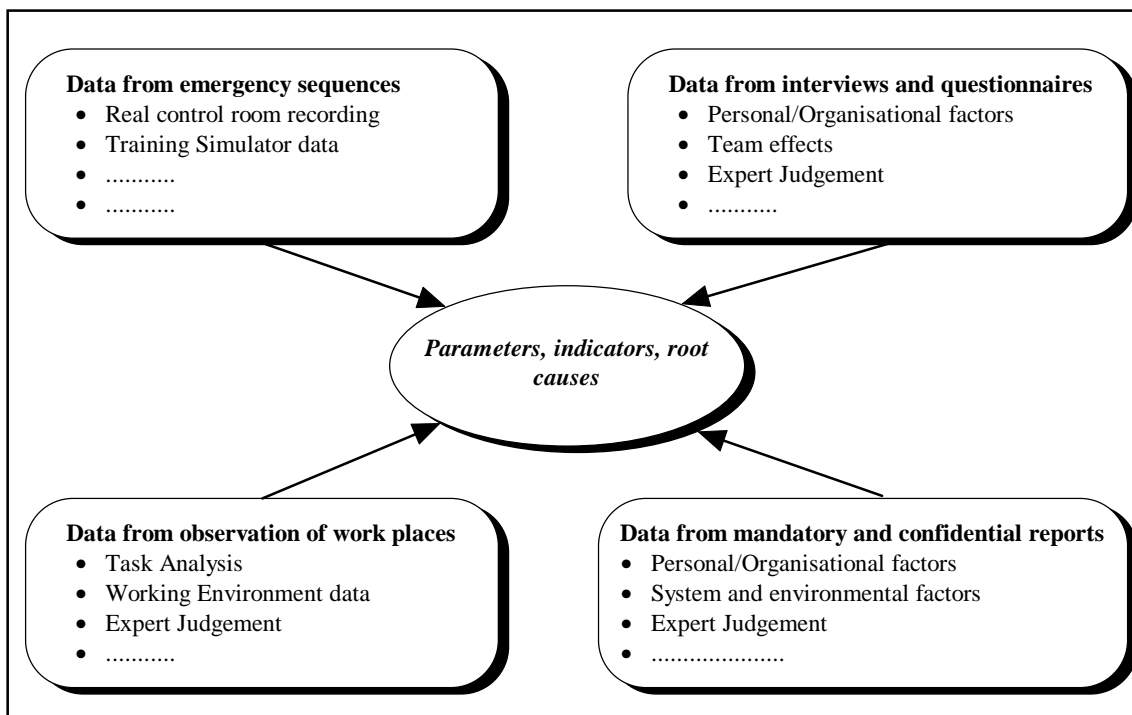


Figure 4-1. Human Reliability Assessment data

The first type of data represent the actual recording of real situations and the results of training sessions, sometime specifically performed with the objective of collecting data about a well defined human-machine situation. These data are very important for defining operator behaviour and response to emergency situations. Some areas in which simulators are useful are (IAEA, 1990; Spurgin and Moieni, 1991):

- detailed analyses of accident scenarios
- generation of non-response (TRC) curves, account for impact of equipment failures on HEPs
- detection of systematic errors (if any) by operator crews
- identification of situations and modes of operation not normally found using more generic HRA techniques
- development of data bases through the use of simulator tests and simulator training
- development and validation of cognitive models with applicability in PSA
- check of opinions of human performance experts

The second type of data are collected in order to develop an understanding of the effects of the working environment and for evaluating the complexity of the tasks of operators. The contribution of the analyst own experience and of expert judgement are crucial in deriving reliability data from these observations.

Data from interviews and questionnaires are very important for the definition of personal and organisational factors affecting behaviour. Once again the expertise of the analyst, in engineering, statistical and behavioural sciences, is essential in deriving risk data from the information collected in this way.

Finally, the existence of data in mandatory and confidential reporting systems on accidents, even in domains different from the one of interest, can be very useful for the definition of human failure probability and root causes of erroneous behaviour.

In order to interpret these data and to develop a database useful for PSA/HRA analysis the use of expert judgement is then essential. In particular, in the area of engineering, it is needed for understanding the physics and material behaviour of the systems, with respect to psychology mainly for the human and social aspects of the interactions, and in connection with statistics for the estimation of data and parameters to be eventually utilised for the HRA.

For nuclear applications, a combination of methods based on simulator data with expert judgement methods remains the most balanced approach to the evaluation of control room performance under accident conditions. Looking to the future, the trend in emerging methods is a significantly increased emphasis on data from operational events. While this data is by nature anecdotal, the general characteristics it presents are highlighting aspects of human reliability that have until now perhaps not been considered sufficiently. See Chapter 8 and Appendix E.

4.4 Chapter References

- /1/ Hirschberg, S., Chakraborty, S. and Kröger, W. (1991), "Human Reliability Data Needs and Potential Contribution of the Halden Project". Workshop Meeting on Human Reliability Data, Halden, Norway, 4th December 1991.
- /2/ IAEA (1990), "Human Error Classification and Data Collection". IAEA TECDOC-538, Report of a Technical Committee Meeting organized by the International Atomic Energy Agency, Vienna, Austria, February 20-24, 1989.
- /3/ Spurgin, A. J. and Moieni, P. (1991), "Interpretation of Simulator Data in the Context of Human Reliability Modeling". Proceedings of the International Conference on Probabilistic Safety Assessment and Management (PSAM), Beverly Hills, California, USA, February 4-7, 1991.
- /4/ Swain, A.D. and Guttman, H.E. (1983), "Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications". NUREG/CR-1278, U.S. Nuclear Regulatory Commission, Washington DC, August 1983.

5. CURRENTLY USED APPROACHES AND THEIR LIMITATIONS

5.1 Introduction to human reliability analysis (HRA) in technical systems

The analysis of human reliability within the framework of a probabilistic risk assessment (PRA) comprises the detection (identification) of safety-relevant man-machine interactions and the calculation (quantification) of the corresponding probabilities [PRA-PG]. This type of analysis offers an opportunity for concrete improvements to the man-machine conditions in a plant. Safety-relevant interactions between man and machine can be predictively identified before they cause or contribute to causing damage (see Figure 5-1). However, this advantage is at the same time a challenge for the analyst. Reliable and practically useful results can only be expected from a high-quality predictive analysis. This quality also depends on the application of relevant insights from psychology and ergonomics and on the application of conclusions drawn from incorrect actions in the past, for example by evaluating operational experience or by means of simulator experiments. In the case of correct application, it will then be possible to realistically distinguish between relatively likely and relatively unlikely damage opportunities.

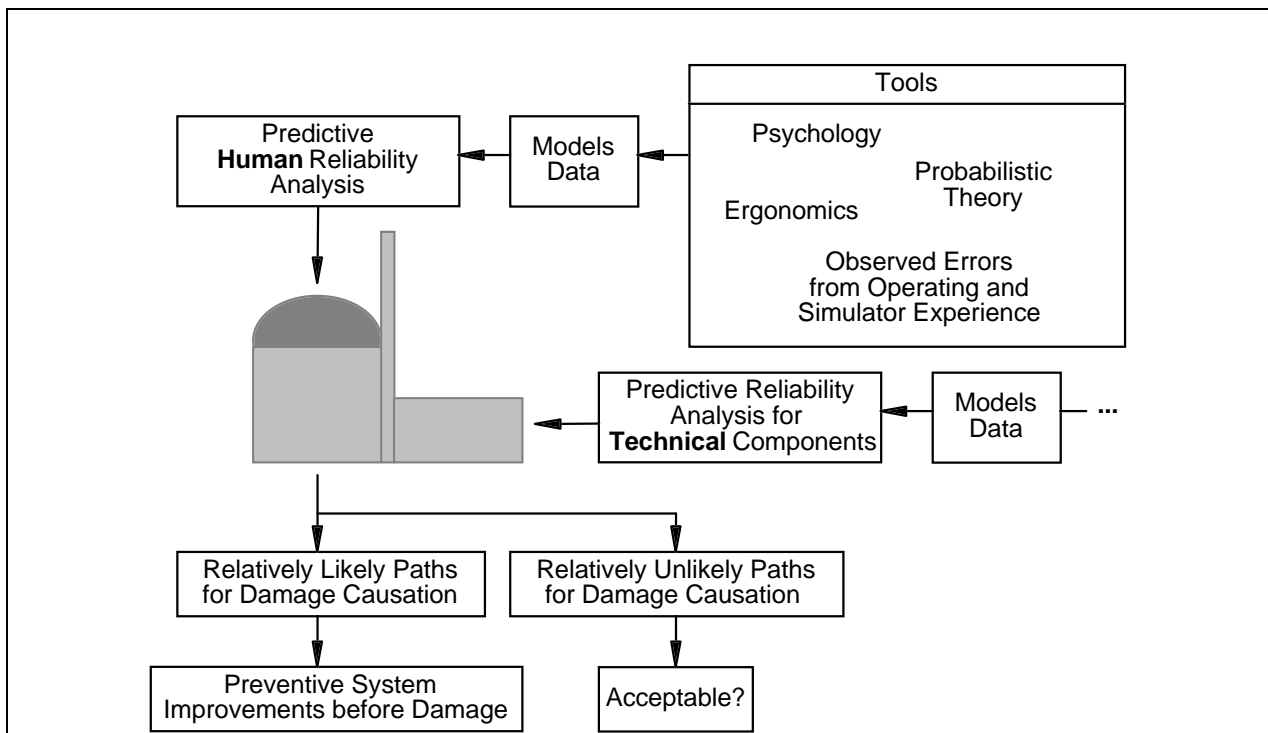


Figure 5-1. Interdisciplinary procedure in predictive human reliability analysis

This report describes the current status of development in the field of human reliability analysis; priority is given to the methods of assessing error probabilities during the course of an accident. Section 5.2 deals with the man-machine interactions, which are normally covered explicitly by a human reliability analysis within a PRA. Section 5.3 outlines the basic procedure in human reliability analysis. The identification of safety-relevant opportunities for human error is the subject of discussion in Section 5.4. Section 5.5 gives a survey of different methods of quantification (for the assessment of error probabilities). Section 5.6

describes and evaluates the established standard method THERP (Technique of Human Error Rate Prediction). Promising new developments are discussed in Sections 5.7 and 5.8. Requirements for an HRA method are outlined in Section 5.9. The report ends with conclusions in Section 5.10. This report is related to part of the work already carried out for [Hennings 95], [Sträter 94b] and [Reer 96].

5.2 Classification of human actions and errors

Human reliability analysis (within a PRA) only comprises a sub-aspect from the total spectrum of human activities in a technical system, for example a nuclear power plant (NPP), i.e. actions by the *operating* personnel *after* commissioning the system. Human errors *prior to* commissioning (design, installation etc.) are assumed to be covered by the reliability characteristics of technical components, see Table 5-1.

Table 5-1. Human activities prior to and after commissioning a plant

<i>Man-machine interaction</i>	<i>Subject of human reliability analysis?</i>
Prior to commissioning: – design – installation – ...	No
After commissioning: – scheduled operation (incl. start-up/shutdown and revision) <ul style="list-style-type: none"> – repair, maintenance, test, calibration – control room activity (monitoring, start-up/shutdown) – accident <ul style="list-style-type: none"> – required and unrequired actions – actions with and without predefined procedure – actions within and beyond design limits 	Yes, with restrictions, see further tables

For the time after commissioning a plant, activities are analysed both in undisturbed normal operation and under accident conditions. The typical results of such activities, which are of interest for a PRA, are listed in Table 5-2.

Table 5-2. PRA-relevant results of human activities

<i>Situation</i>	<i>Man-machine interaction</i>		<i>Subject of human reliability analysis?</i>
	<i>positive</i>	<i>negative</i>	
normal	malfunction of a component <u>detected</u>	malfunction of a component <u>overlooked</u>	YES, with restrictions according to Table 5-3
	tested / serviced / repaired component <u>left in proper condition</u>	<u>malfunction</u> of a component <u>brought about</u>	
abnormal (accident)	required manual action for accident control <u>successful</u>	required manual action for accident control <u>fails</u>	YES, with restriction according to Table 5-9
	unrequired and dangerous intervention in automatic system responses for accident control <u>omitted</u>	unrequired and dangerous intervention in accident control <u>made</u>	SELDOM (requirement of methods)

5.2.1 Activities in normal situations (prior to the occurrence of an accident)

5.2.1.1 Latent errors contributing to the unavailability of safety-relevant systems

Not all activities in the absence of an accident are the subject of human reliability analysis. Normally, only those activities are analysed in which errors (e.g. faulty maintenance) can occur leading to failure of a component included in the fault trees. These errors correspond to the man-machine interaction classified as type 1 errors in [SHARP] or type A in [IAEA 89]. They thus occur primarily during repair and maintenance work. Their analysis is not required if the contribution of such activities as, for example, maintenance, is already covered by the failure statistics of the component concerned. However, corresponding evidence is not always easy to furnish, especially if

- the procedure belonging to the activity, generally a written instruction for action, is new,
- the activity is very plant-specific or
- failure of the component concerned is very seldom.

A contribution to the failure of a fault tree component, quantified by human reliability analysis, is generally a latent (hidden) error (cf. [Reason 90], Chapter 7; see also Table 5-3): a system or component is in an improper standby condition, possibly for a prolonged period of time. Table 5-4 contains two examples of latent errors from operating experience with nuclear plants. In the most favourable case (Gundremmingen) such an error is detected during later inspection or checking, in the most unfavourable case (TMI) only during an accident when the system affected is urgently needed.

5.2.1.2 Active errors initiating accidents

Human reliability analysis rarely deals explicitly with errors initiating accidents (Table 5-3). Table 5-5 contains two examples from nuclear operating experience. Such errors are practically always assumed to be covered by the frequency statistics of accident-initiating events. Rough estimates – as, for example, in [KFA 81] for errors leading to a loss of the main heat sink – are the exception. Based on recent insights [EPS 1300] into the possible risk relevance of zero power level conditions, there is increasing interest in a precise analysis of accident-initiating errors. Zero power level conditions in nuclear power plants require a large number of additional manual actions which do not occur during power operation. Recent American studies confirm that most (30 out of 39) of the errors recorded in zero and partial power operation resulted in the initiation of an accident ([Barriere 94], page 2-8).

Table 5-3. Errors covered by human reliability analysis (HRA), broken down according to plant state and degradation of plant safety

<i>Degradation of plant safety</i>	<i>Plant state</i>		
	<i>zero power, revision</i>	<i>partial power, start-up/shutdown</i>	<i>full power operation</i>
latent error (after maintenance, repair, test or calibration or control room activity): component / system undetected unavailable	covered by human reliability analysis for relevant fault tree inputs		
active error (during maintenance, repair, test or calibration or control room activity): accident is initiated	not covered by HRA with a few exceptions	not covered by HRA with a few exceptions, e.g. [KFA 81]	

Table 5-4. Examples of latent errors caused by the operating personnel

<i>Error</i>	<i>Degradation of plant safety</i>
NPP Gundremmingen, 1975 [BMI 77]: primary isolation valve closed during calibration and not re-opened	measuring channel group for differential pressure unavailable
NPP TMI-2 (Harrisburg), 1979 [Kemeny 79]: block valves closed during maintenance and not re-opened	auxiliary feedwater system unavailable

Table 5.5. Examples of accident-initiating (active) errors of the operating personnel

<i>Error</i>	<i>Degradation of plant safety</i>
NPP Obrigheim, 1972 [IRS 77]: drain valve inadvertently opened when purging the pressurizer relief tank	loss of coolant
NPP Neckarwestheim, 1975 [Smidt 79]: demineralized-water feeding stopped too late during start-up	unintended criticality

5.2.2 Activities in abnormal situations (during accidents)

5.2.2.1 Required and unrequired actions

The human reliability analysis of activities during accidents is generally confined to manual actions required for accident control. In exceptional cases, unrequired and hazardous interventions in automatic system responses are also studied, see Table 5-2. Evaluations of operational experience specify an amount of 22 % for unrequired and unsuitable actions; see Figure 5-2.

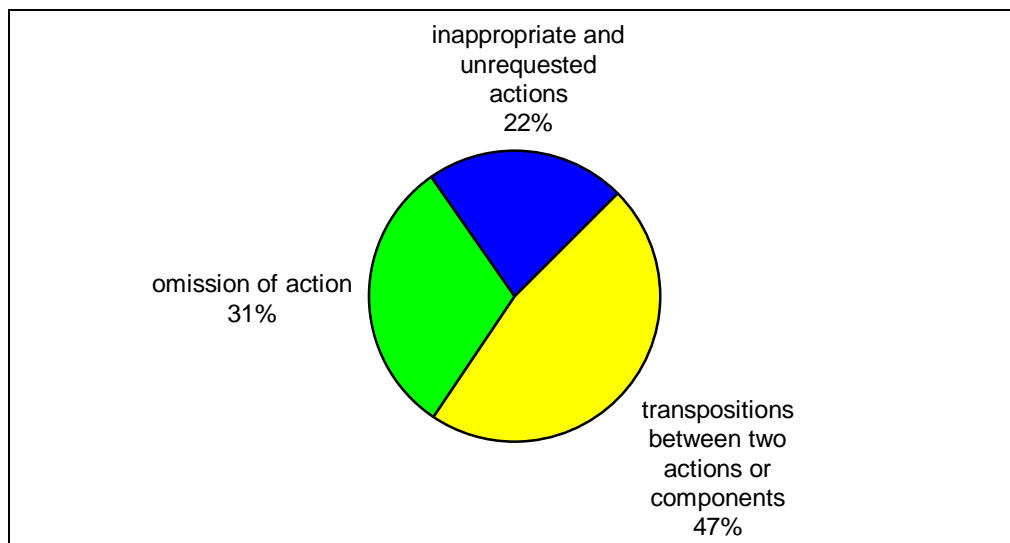


Figure 5-2: Percentages of different types of human errors for 51 events in nuclear power plants [Ghertman 85]

Against this background, the unrequired shutdown of safety systems is also analysed for French nuclear power plants [EPS 900], [EPS 1300]. Such actions are also referred to in the PSA literature as unsuitable initiatives ([EPS 900], page 73), extraneous acts ([Swain 83], page 4-9) or mistakes that aggravate situations ([SHARP], page 2-6). Table 5-6 shows that unrequired actions can have both a positive and a negative influence on the course of an accident. It cannot therefore be stated from the outset whether the risk is underestimated or overestimated if a PSA neglects such actions. Table 5-7 contains examples of required manual actions during accidents, whose failure is studied within the scope of a PRA. However, their requirement (with a view to avoiding core damage) does not always result exclusively from the accident-initiating event. In many cases, an action (e.g. isolation of a defective steam generator) is only required in combination with a further malfunction (e.g. failure of the water feed system replenishing any water lost through steam generator leakage) in order to prevent core damage. In rare cases, a manual action is required exclusively due to the initiating event; e.g. shutdown after a small leak in the main coolant pipe of older plants in which the shutdown process is not automated.

Table 5-6. Examples of unrequired operator actions from operational experience with nuclear power plants

<i>Action</i>	<i>Unrequired because ...</i>	<i>Course of action</i>	<i>Influence on accident sequence</i>
NPP Obrigheim, 1972, after loss of coolant through open drain valve (TE1A10): bridging the torque limit switch of TE1A10 [IRS 72]	...automatic safety injection would have controlled the loss of coolant [Hoffmeister 74]	positive: bridging was successful, so that valve (TE1A10) could be closed	positive: loss of coolant stopped
NPP Davis Besse, 1985, after failure of the main feedwater pumps: manual start of emergency feedwater pumps [NUREG-1154]	...the emergency feedwater pumps would have started automatically later on	negative: starting fails, steam generator inadvertently isolated	negative: failure of both emergency feedwater pumps
NPP TMI-2, 1979, after loss of coolant through open pressurizer relief valve: manual disconnection of the high pressure injection (HPI) pumps [Kemeny 79]	...automatic HPI would have controlled the loss of coolant	positive: disconnection of HPI pumps successful	negative: core damage

Table 5-7. Important operator actions included in the DRS-B accident sequence diagrams (event trees) and quantified (after [DRS-B], Section 5.2)

<i>Required action</i>	<i>Accident-initiating event</i>			
	<i>small leak in a main coolant pipe</i>	<i>leak in a heating tube of a steam generator</i>	<i>failure of main feedwater system</i>	<i>failure of the main heat sink</i>
plant shutdown	X	X		
start of emergency system		X	X	X
isolation of defective steam generator		X		
restart of volume control system		X		

5.2.2.2 *Actions with and without predefined procedures*

Not every action (or sequence of actions) suitable for controlling an accident sequence is predefined by a written instruction. Actions without predefined procedure are referred to by [SHARP] (page 2-6) as type 5 actions: »By improvising, plant personnel can restore and operate initially unavailable equipment to terminate an accident«. Such actions are basically repairs (in the widest sense) to failed components. The specific course of a repair depends on the failure cause of the component concerned, in one case only a fuse has to be replaced or an actuator pressed (e.g. manual start on failure of automatic starting device), in another the whole component must be replaced. Since there are often several failure causes, some of them unknown, *concrete* repair actions are seldom laid down in procedures and are therefore not *explicitly* included in risk studies. Several risk studies [KFA 81], [KFA 84] for plants, in which there is much time (more than about 5 h) available for repairing a failed component during an accident, explicitly include repairs. The error probability is estimated by evaluating a repair statistics documenting the repair times required for failures which previously occurred on the component concerned. Methodological approaches to evaluation and critical comments on the informative value of such statistics can be found in the literature [Reer 94a].

The extent to which manual actions described by procedures are included depends on the scope of analysis defined in the relevant PRA. A basic criterion is the question of the extent to which recovering interventions by the operating personnel are analysed in beyond design basis accident situations (“accident severity”); this will be discussed in the following section.

5.2.2.3 *Actions within and beyond design*

Actions for accident control are classified into two groups in Germany (cf. [Roth-Seefried]; [DRS-B]):

1. actions within design basis accident control, e.g. start of emergency system after failure of steam generator main and auxiliary feeding,
2. emergency actions beyond design basis accident control, e.g. steam generator feeding by a mobile fire fighting pump after failure of main feeding, auxiliary feeding and emergency systems.

The first group is related to the *planned* operating and safety systems for accident control under the licensing procedure, so that the term “planned actions” (in German: “geplante Handlungen”) is also used. Most of these actions are described in the operating procedures. There are, however, also actions within design basis accident control, which are not contained in the operating procedures, for example the manual start of a system upon failure of the automatic starting device.

There are often still safety margins available even for a failure of design basis system functions. In these cases, the accident can still be controlled by emergency actions, although this is already an *unplanned accident* according to licensing design criteria. For this reason, the terms “unplanned actions” (in German: “ungeplante Handlungen”) and “accident management” are also used in this connection. The action sequences are described in emergency operation procedures, inasmuch as they can be predefined in detail. In international usage, such actions are subsumed under the term “recovery” (“... all human interactions that are not initially integrated into plant logic models”) [Wakefield 92] (page 2-2).

Although emergency actions do not fall within the scope of design basis accident control, considerations have recently been under way to the effect of including such actions in the regular simulator training programme [Hoffmann 89]. According to [Swain 83] (page 12-23), a reliable diagnosis (i.e. the timely recognition of the need for an emergency action) depends on such training. The emergency exercises recently introduced in many nuclear power plants [Hansmann 89], [Gautschi 89] have a direct influence on diagnosis reliability. Exercises of the required manual actions reduce their time requirement, so that more time is available for diagnosis and the operators can analyse the system state under less time pressure. Knowledge of the training level in the plant under consideration is thus of great significance for human reliability analysis.

The question of the extent to which such emergency actions have been included depends on the objectives of the study. In many studies (e.g. [EPS 1300]) no distinction is made (as, for example, in Germany) between emergency actions and actions within design basis accident control. A relevant evaluation of such studies therefore requires a determination of the extent to which accident sequences, for which manual operator actions for accident control were still included, have exceeded the scope of design basis accidents.

5.2.3 *Error causation categories*

Various types of errors can contribute to the failure of an action or sequence of actions. The categories of incorrect human outputs in [Swain 83] provide rough insights into error causation. Deeper insights can be found in classifications based on modern psychological error research. In this context, the error type classifications in [Reason 90] and [Hacker 86] have gained acceptance.

5.2.3.1 *Errors of omission, errors of commission*

Table 5-8 outlines a classification system taken from the final version of the THERP handbook [Swain 83]. Two main groups are distinguished: error of omission (EOM): failure to perform a task (or action) and error of commission (ECOM): incorrect performance of a task (or action).

It should be noted that any classification according to these groups (defined in [Swain 83]) depends on the definition of the underlying task (or action). Given a definition including a very detailed specification of the respective task, almost every error can be classified as an EOM. For example, during the Davis-Besse incident (see Table 5-6), an operator pushed the wrong controls labelled “SG 1-1 LOW STM PRESS” and “SG 1-2 LOW STM PRESS” while attempting to actuate the auxiliary feedwater system (AFWS), he should have pushed the controls labelled “SG 1-1 LOW WTR LVL” and “SG 1-2 LOW WTR LVL”. This is an ECOM according to a task defined as ACTUATE AFWS. However, according to a task defined as PUSH CONTROLS “SG 1-1 LOW WTR LVL” and “SG 1-2 LOW WTR LVL”, this is an EOM.

To avoid such pitfalls, an ECOM should be defined as an incorrect performance of a task, given that the task is initiated. Moreover, it should be noted that the boundaries between the ECOM categories are fluid. For example, an ERROR OF SEQUENCE could also be described by the attributes TOO EARLY or TOO LATE. Nevertheless, the classification system in Table 5-8 provides valuable insights into error causation because:

- it demands that an analyst describes the task of an operator, and
- it gives an overview of the principal reasons why a task could fail.

Table 5-8. Categories of incorrect human outputs. “Any of these incorrect human outputs may be the result of other human errors: an error of interpretation of a pattern of signals, a misreading of a display, a misprint in an emergency operating procedure, etc. In an HRA, the incorrect human outputs and human errors leading to these incorrect outputs must be analysed “ ([Swain 83], page 2-16)

<ul style="list-style-type: none"> – – 	<ul style="list-style-type: none"> Errors of Omission <ul style="list-style-type: none"> • Omits entire task • Omits a step in a task Errors of Commission <ul style="list-style-type: none"> • Selection error: <ul style="list-style-type: none"> - Selects wrong control - Mispositions control (includes reversal errors, improperly made connections, etc.) - Issues wrong command or information (via voice or writing) • Error of sequence • Time error: <ul style="list-style-type: none"> - too early - too late • Qualitative error: <ul style="list-style-type: none"> – too much – too little
--	---

5.2.3.2 *Slips, mistakes, violations*

Reason's classification (Table 5-9) of errors - or (more general) unsafe acts - uses basic elements of the cognitive behaviour classification in [Rasmussen 79] into the following levels:

- skill-based (stored patterns of pre-programmed instructions are called from memory),
- rule-based (familiar problems are tackled with the aid of IF-THEN rules), and
- knowledge-based (novel problems are tackled by analytical approach).

This classification forms the basis in the USA for categorising human error types in events occurring in nuclear power plants (cf. [Barriere 94], Appendix A). With respect to these properties of human cognitive behaviour, Reason distinguishes two main groups of errors:

- A. Wrong execution of a plan due to attentional failure (e.g. overlooking something) or memory failure (e.g. forgetting something): error types 1 and 2 in Table 5-9. These errors are summarised as “skill-based errors” and specified as “slips” and “lapses”.
- B. Erroneous or deliberate decision for a wrong plan: error types 3 to 7 in Table 5-9.

Table 5-9 suggests that the current methods of human reliability analysis exhibit weak points in the consideration of decisions for wrong plans. Erroneously wrong decisions are frequently only quantifiable at the rule-based level, i.e. as errors in the application of IF-THEN rules: if system state ZX, then plan PY ([Reason 90], page 43).

Knowledge-based mistakes, for example selective perception directed to a misleading signal, are to be expected in novel situations involving limited resources in conjunction with knowledge deficiencies for problem solution ([Reason 90], page 43). All methods of human reliability analysis exhibit deficiencies in the quantification of such mistakes, and models are used which greatly simplify reality (e.g. the diagnosis model in [Swain 83]).

Deliberately wrong decisions (violations) form a special category. A violation of type 5 or 6 (Table 5-9) is not necessarily an error in the sense that the goal envisaged by the operator is not achieved, but a deviation from the procedures (or policies) defined by designers, managers and authorities to ensure plant safety ([Reason 90], page 195).

The operator thus runs the risk in the event of an unsafe deviation from such prescribed procedures that an undesirable system state occurs with increased probability. Routine violations, such as subsequently filling in a checklist, are basically covered by normal human reliability analysis (e.g. [Swain 93], Table 20-6). There are, however, deficiencies in the quantification of violations in exceptional situations, e.g. during an accident. Some methods ([Hall 82], page 26; [Moieni 94], page 54) generally include such action tendencies by means of a so-called “reluctance factor” by which the error probability of a task involving economic losses is increased (e.g. depressurization of the primary circuit after total failure of steam generator feeding, which results in a prolonged plant outage).

Intended sabotage is not a subject of the HRA methods described here. It is always assumed that the operator is willing and tries to avert damage from the system (be it by a wrong or correct plan).

Table 5-9. Error types included in human reliability analysis

<i>Error types after [Reason 90] which can contribute to failure of an action (or of a sequence of actions)</i>	<i>Explicitly included in human reliability analysis?</i>	
action is <i>not</i> performed as planned (skill-based error)	(1) attentional failure (slip)	basically yes, point-type deficiencies *3)
	(2) memory failure (lapse)	
action is performed as planned, <i>plan is erroneously wrong</i> *1) (mistake)	(3) rule-based mistake	in part, strongly simplifying models with frequently estimated input parameters *3)
	(4) knowledge-based mistake	
action is performed as planned, <i>plan is deliberately wrong</i> *2) (violation)	(5) routine (frequently occurring) violation	basically yes, point-type deficiencies *3)
	(6) exceptional (seldom occurring) violation	in part, strongly simplifying models with frequently estimated input parameters *3)
	(7) act of sabotage	no

*1) Unsuitable for achieving the goal envisaged by the operator.

*2) Deviating from rules (procedures or policies). Note, in contrast to other authors, Reason does not define such deviation as a criterion for an error; cf. [Dougherty 94] and [Reason 94].

*3) The assessments concerning a consideration of these error types are vague because the boundaries – e.g. between (4) and (2) – are not always very sharp.

The error types 1 to 4 are specified by 35 sub-categories outlined in [Reason 90] (pages 68-95).

In practical HRA application, it is difficult to distinguish between the categories in Table 5-9. As an example, the HCR model [Hannaman 84] differentiates between everyday actions for skill-based behaviour, actions with procedures for rule-based behaviour, and actions without procedures for knowledge-based behaviour. Nevertheless, the relation of cognitive behaviour to real objects is not a 1:1 representation and categorising into these cognitive levels is more difficult than expected. Even the application of available IF-THEN rules frequently requires knowledge-based performance, for example if the system state must be determined from several interacting process parameters. According to evaluations in French nuclear power plants, even the application of a written procedure contains knowledge-based performance elements:

»The operator does not apply the procedure mechanically. ...Every time an operator reads a procedure he *necessarily* interprets or “re-thinks” it« ([Mosneron 92], page 631).

Considering this example, HRA has to recognise that the cognitive behaviour of human beings is not related to external properties of the situation (i.e., procedures do not always involve rule-based behaviour). Human beings are able to accommodate themselves to different situations. They learn behaviour and become skilled if they do an action several times. If an action is performed at least regularly and often, it may become highly skilled and subconscious. This enables the human being to reduce information load by performing some cognitive activity on the subconscious level and is to a wide extent independent of the observable complexity of the action. Hence, an application of a bad procedure may be a knowledge-based decision error as well as a skill-based error of habit. This depends on the situation (external cues of information) as well as on the experience and habit of the operator.

The basis for this accommodative performance is the recognise-act cycle of human information processing. If the recognition of a situational pattern does not imply an abnormality (e.g., a new situation) in the human brain, then the human will act on the subconscious or low conscious level (e.g., walking until we stumble). If the recognition does imply an abnormality then conscious processing will be involved for this abnormality. This so-called 'cognitive dissonance' is a mismatch of learned behaviour and actual situation.

This means that the level of cognitive behaviour is directly related to the amount of cognitive dissonance which is implied by the difference of the situation from a learned behaviour. The distinction of the level of cognitive behaviour is therefore more a continuum measure than a dichotomy and has to consider both the learned behaviour of the operator and the involved tasks (cf. [Wickens 84]). This aspect may be summarised in Table 5-10. [Sträter 95b] provides an elaborated discussion of cognitive dissonance, underlying cognitive dimensions and resulting cognitive habits.

The table indicates that the recognition of the cognitive level by the operator is normally not distinguished in HRA. The underlying assumption for this is that operators are normally well trained (i.e., the assumption of the log-normal distribution) and differences in the operator's recognition of the cognitive level are modelled by PSFs like training and qualification.

In summary, it should be noted that the deficiencies (indicated in Table 5-9) considering decision-based errors in action planning do not necessarily suggest that the current HRA approach leads to unrealistic results. Not every pursuit of a wrong plan can be classified as a mistake or violation. Skill-based errors (e.g. overlooking or disregarding something) can also cause an operator to follow a wrong plan. A certain percentage of decision-based errors is therefore explicitly quantifiable by probabilities of errors in skill-based performance. The method proposed in [Moieni 94] for assessing the probability of misinterpreting a procedure step containing a diagnosis logic is based e.g. on the THERP data for rule-based actions; see

Figure A-6 in Appendix A.

Moreover, it cannot be ruled out that the error probabilities used in human reliability analysis also implicitly include contributions beyond the skill-based level. A certain percentage of errors from which an error probability was calculated could be attributable, for example, to knowledge-based mistakes (such as excessive confidence in a misleading information perceived first).

Table 5-10. Coherence of cognitive level and situational factors and involved error types

<i>Operators' recognize-act cycle</i> <i>situation</i>	<i>skill</i> <i>(no dissonance)</i>	<i>rule</i> <i>(dissonance)</i>	<i>knowledge</i> <i>(strong dissonance)</i>
skill (regular and frequent everyday-action)	skill-based slip	rule-based slip	knowledge-based slip
rule (regular and frequent action with procedure)	routine violation	rule-based mistake	attentional failure
knowledge (not regular and unusual without procedure)	exceptional violation	memory failure	knowledge-based mistake

5.2.3.3 Lack of information, non-use of information, incorrect use of information

The classification system of Hacker ([Hacker 80], revised version in [Hacker 86]) emphasises the underlying information deficiency of an error. Table 5-11 summarises the related categories. Similar to the categories EOM and ECOM in [Swain 83] (see Table 5-8 in this report), Hacker distinguishes between NONUSE and INCORRECT USE of available information. There are also similarities (e.g., B.3 violation such as ignoring available information) to categories used in [Reason 90] (see Table 5-9 in this report).

In addition to the error classification systems outlined above, Hacker's system offers an important category: lack of useable information. This category allows for an advanced view of human error as man-machine interaction: human deficiencies and deficiencies of the technical systems contribute to the causation of an error. For example, it is evident that a technical deficiency (incorrect signal of valve state) was an important cause of a significant error (cutting out HPI pumps) during the TMI incident (1979); see Table 5-3.

Table 5-11. Main headings of categories for information deficiencies as causes for human errors. Translated from [Hacker 86] (page 435). The categories C.1, C.2 and C.4 are specified by eight sub-categories, see [Hacker 86], Section 10.2 for details.

<p>(A) LACK OF USEABLE INFORMATION</p> <p>(B) NONUSE OF AVAILABLE INFORMATION</p> <p> (B.1) Overlooking</p> <p> (B.2) Forgetting</p> <p> (B.3) Ignoring (violation)</p> <p> (B.4) Information reduction (stereotype response without verifying check)</p> <p> (B.5) Deficiencies in information processing due to time restrictions or mental capacities</p> <p>(C) INCORRECT USE OF AVAILABLE INFORMATION</p> <p> (C.1) Incorrect use in orientation</p> <p> (C.2) Incomplete use in goal formation</p> <p> (C.3) Formation of an incorrect plan to act</p> <p> (C.4) Mismatching recall of a correct plan to act</p>

5.3 Basic procedure in human reliability analysis

5.3.1 HRA steps

Essentially, many of the proposed sequences of HRA steps have equal contents. However, according to the starting point, two different concepts appear worth distinguishing:

- starting with a human-related event (identified by fault tree analysis) [PRA-PG] or system failure of interest [Swain 83], or
- starting with a certain type of human interaction ([IAEA 89], [SHARP]).

5.3.1.1 PRA Procedures Guide, THERP

Table 5-12 outlines the basic procedure recommended in the American PRA Procedures Guide ([PRA-PG], Chapter 4) for human reliability analysis. The procedure comprises 12 steps, which are very much tailored to the THERP method (Technique of Human Error Rate Prediction) described in Swain's handbook [Swain 80], [Swain 83]. Since other methods are also discussed here, this section concentrates on a simplifying classification of the essential steps (1 to 10) into two groups:

- error identification, and
- error quantification.

Note that in [Swain 83] (Fig. 5-6) a more detailed classification is given: familiarisation (steps 1-2), qualitative assessment (3-5), quantitative assessment (6-10), and incorporation (11-12).

Table 5-12. Rough classification of steps recommended for human reliability analysis in the PRA Guide [PRA-PG]. The steps are to apply to each “human-related event” identified by the fault tree analyst.

<i>Step</i>	<i>Level of analysis</i>	
	<i>error identification</i>	<i>error quantification</i>
1) plant visit	X	
2) review information from fault tree analyst	X	
3) talk-through	X	
4) task analysis	X	X
5) develop HRA event trees		X
6) assign human error probabilities (HEPs)		X
7) estimate the relative effects of performance shaping factors (PSFs)		X
8) assess dependence		X
9) determine success and failure probabilities		X
10) determine the effects of recovery factors		X
11) perform a sensitivity analysis, if warranted		
12) supply information to the fault tree analyst		

Section 5.4 (corresponding to steps 1 to 4 in Table 5-12) gives a brief survey of the methods of predictive identification of safety-relevant errors. Section 5.5 (steps 4 to 10) contains a rough evaluation of various probabilistic error quantification methods.

The classification selected here emphasises the central significance of task analysis for the evaluation of human reliability. It comprises ([PRA-PG], Chapter 4) a decomposition into sub-tasks (sub-actions) and the identification of safety-relevant error opportunities.

Step 4 referred to as task analysis is closely connected with the other steps. A task analysis is based on the plant visit, review of human-related events identified by fault tree analyses and discussion of the procedure steps of the manual action to be analysed together with the competent operators. The most important factors to be included are compiled in [DRS-B] (pp. 399-400):

- task description, i.e. a brief description of the tasks to be fulfilled by the personnel,
- location, i.e. specification of the location at which the action is performed,
- time after the occurrence of the accident, i.e. time at which the action must be initiated at the latest,
- duration of the action, i.e. average time required for performing the action,
- written instructions or aids, i.e. stating whether and, if any, what written instructions or working documents are available to the personnel for performing the action,
- possibility of identifying the need for manual actions, i.e. what means of information are available to the personnel indicating the need for performing the action (only basic information sources are included),

- substitution for automatic action, i.e. stating whether the action described must be performed in the absence of an action otherwise initiated automatically,
- cause, i.e. stating the reasons requiring the action,
- action feedback, i.e. stating the means of information which provide the personnel with a feedback of successful task performance,
- relation to other actions, i.e. information about functional relations with other actions.«

This provides important information for a realistic quantification in steps 5 to 10 (Table 5-12); e.g.

- functional relations between sub-actions for realistic modelling in step 5, or
- conditions aggravating or facilitating task performance for a realistic matching of HEPs and conditions in the plant to be analysed (steps 6 and 7).

The evaluation of individual operator performances is not the subject of a task analysis.

5.3.1.2 IAEA, SHARP

The THERP handbook ([Swain 83], Chapter 4) had set up 10 steps for man-machine system analysis (MMSA). These steps can be arranged as a subset of the HRA process described in [IAEA 89], which is similar to the approach in [SHARP]. If the MMSA steps in [Swain 83] are taken as sub-steps of the IAEA structure, the HRA process is divided into the following steps which have to be repeated to optimise MMS reliability. Note that steps 9 and 10 of Swain are excluded because they are relevant for the design process but not for the HRA process.

1. Definition: Analysis of the different types of human actions (Types A-C in [IAEA 89], corresponding to types 1-5 in [SHARP]).
2. Screening: Identify the human interactions that are significant for the operation and safety of the plant.
 1. Describe the system goals and functions of interest.
In this step the points of interaction have to be identified.
3. Qualitative Analysis: Detailed description of the important human interactions and definition of the key influences.
 2. Describe the situational characteristics.
In this step the external PSFs have to be identified.
 3. Describe the characteristics required of the personnel.
In this step human-related PSFs (internal PSFs) have to be identified.

4. Representation: Modelling of human interactions in logic structures.
 4. Describe the jobs and tasks performed by the personnel.
In this step a task analysis has to be performed. Note that a task analysis for HRA purposes is generally not performed in as much detail as for ergonomic design.
5. Impact Integration: Exploration of the impact of significant human actions.
 5. Analyse the jobs and tasks to identify error-likely situations and other problems.
In this step the task has to be brought into relation to the situational and personnel characteristics.
6. Quantification: Assign probabilities for the interactions.
 6. Estimate the likelihood of each potential error.
In this step the dependence of sub-tasks has to be estimated.
 7. Estimate the likelihood that each error will be undetected or uncorrected.
In this step, recoveries and consequences have to be estimated.
 8. Estimate the consequences of each undetected or uncorrected error.
In this step a sensitivity analysis has to be performed.
7. Documentation: Making the analysis traceable, understandable, and reproducible.

Summarising, one can find two different requirements concerning the type of data and the level of data. The first requirement is related to the qualitative type of HR data while the second one is related to the quantitative type of HR data. They are summarised in Table 5-12.

5.3.2 Data acquisition for HRA

5.3.2.1 Requirements of the HRA methods

In [IAEA 90], the requirements of the different methods have been summarised. In the document one can find the following requirements:

The HCR model [Hannaman 84] needs data about

- the working media or the man-machine interface
- influence factors e.g. stress, skill level, fatigue, environment and organisational effects
- the time available for diagnosis and correct execution of a task (time window for action).

The SLIM approach [Embrey 84] needs data about

- the description of the task and actions
- influence factors e.g. stress, skill, fatigue, environment and organisational effects
- the number of demands for two tasks for calibration of the quantitative estimations.

The THERP method [Swain 83] needs data about

- the description of the task and the actions
- influence factors e.g. stress, skill level, fatigue, environment and organisational effects
- recovery factors for the different tasks
- the working media or the man-machine interface
- the persons or teams which have to perform the task
- the time available for diagnosis and correct execution of a task (time window for action)
- the available procedures
- the dependence of the different tasks and influence factors.

The ASEP method [Swain 87] needs data about

- the technical system
- the available procedures
- the time available for diagnosis and correct execution of a task (time window for action)
- recovery factors for the different tasks
- the dependence of the different tasks and influence factors
- the error type.

5.3.2.2 *A framework to acquire HRA data*

The approach presented in the following is a bottom-up approach that is able to provide data for different current HRA methods (for discussion see [Sträter 94a]). It has the advantage that the data obtained can be used for most of the present HRA purposes. The framework is outlined in Figure 5-3.

The framework is based on the MMS (man-machine system) approach (see [Bubb 93]). A MMS can be described in the following way: A task description will be evaluated by an operator. The operator processes the task with respect to the actual system state (feedback loop). By performing an action via the control elements he changes the system behaviour. The system change is then displayed and reported to the operator. The whole MMS is embedded in an organisation structure and an environment. The organisation is related to shift or revision plans or general strategies to improve safety for instance. Here everything can be noted, which is not directly related to the work of the person. In some situations the operator also has to report the task transaction to the supervisor or has to document his transaction in a test protocol or written procedure.

To extend this MMS approach to an error modelling approach it is necessary to add influences to the different stages of the MMS as indicated by the dotted arrows in Figure 5-3. These influences are known as performance shaping factors (PSFs). Internal PSFs can be found in the human-related arrows and external PSFs in the other arrows.

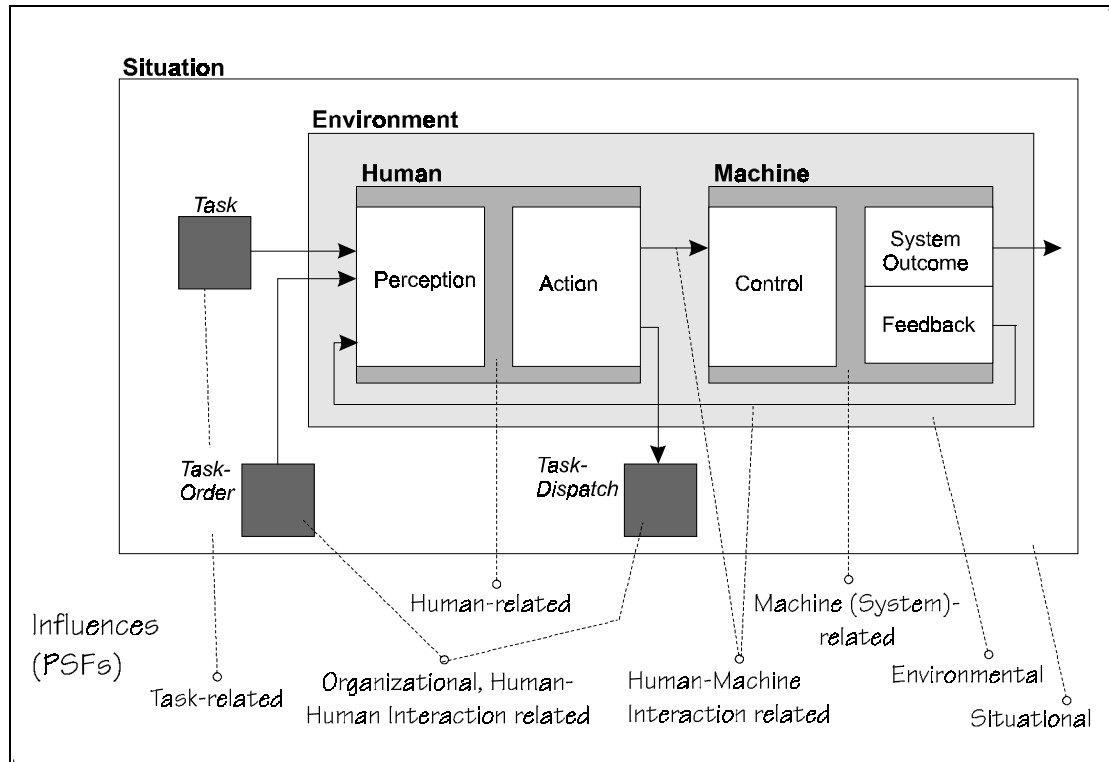


Figure 5-3: A framework to collect human reliability data for various HRA methods.

From the framework, nine questions can be derived as outlined in Table 5-13. Answering these questions leads to complete information about an event, if the following properties are noted according to the different stages of the MMS: factual information, error taxonomies (like EOM/ECOM in Table 5-8), and influences. If a taxonomy is used to fill up the framework the acquired data become highly reliable. For more details on this framework see [Sträter 95a].

With this scheme it is possible to acquire data for HRA purposes from events or simulator data or other types of low level information (e.g., experiments). The framework introduced in this section is able to provide most of the information which is necessary for the HRA methods. How to use this information is described in the methods.

Table 5-13. Nine questions on data for HRA.

1	Task	What was the task of the person ?
2	Person	Who was the person involved in the event ?
3	Action	What has the person done ?
4	Feedback	What were the information sources to check what has to be done ?
5	System	What part of the system has been manipulated ?
6	Environment	Where has the event happened ?
7	Transaction	What had the person to do to inform others about his work ?
8	Organisation	What are the contributing organisational factors ?
9	Time and Duration	When did the event happen and how long did it last ?

5.4 Identification of safety-relevant errors

5.4.1 Analytical expertise

The methods of hazard and error identification summarised in [Madjar 93], [ESCIS 86] and [Whalley 89] can be used, in principle, for the identification of safety-relevant error opportunities contributing to a system failure of interest or to the failure of an action (or sequence of actions). Error identification in human reliability analysis within a PRA is generally inductive and deductive without any rigid scheme being given (»no hard-and-fast rules« [Swain 83], page 4-9). The prediction of safety-relevant error opportunities is mainly based on the analyst's analytical expertise. It is thus a highly knowledge-based activity. Important preliminary information on safety-relevant operator errors is provided by the deductive approach of fault tree analysis. The systematic search for failure causes of a system function enables the detection of human errors related to these failure causes; see Figure 5-4 and Table 5-2.

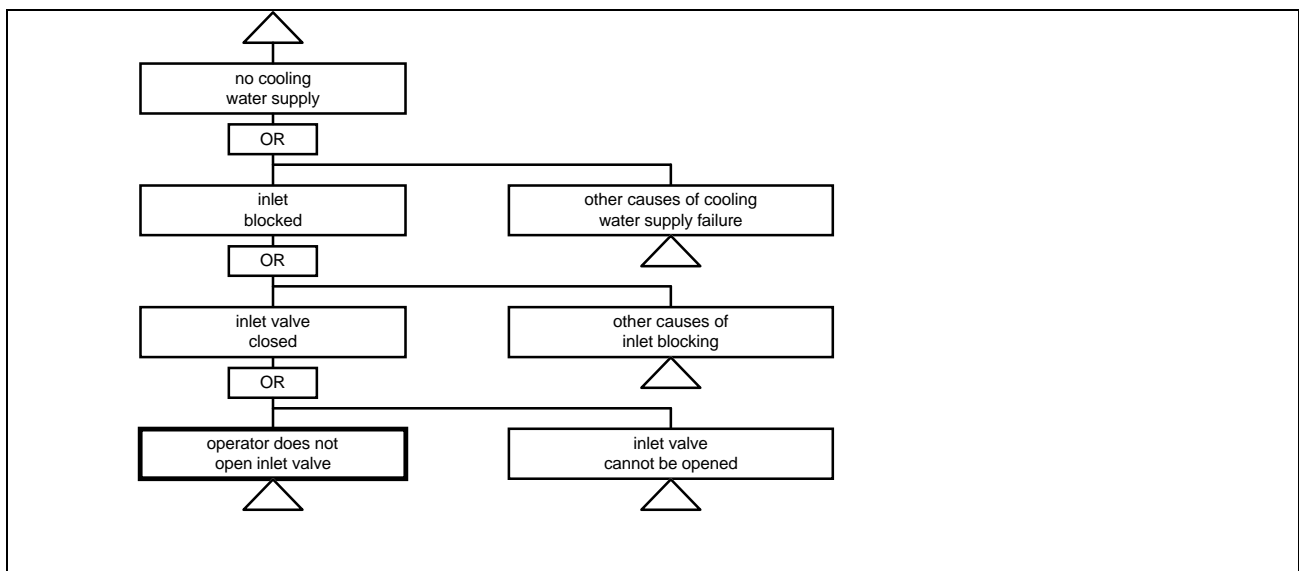


Figure 5-4. Deductive procedure for the identification of operator errors contributing to system failures

As a rule, the review of available fault trees (step 2 in Table 5-12) is not sufficient to ensure a high degree of completeness in error identification. For this reason, the single sub-tasks, into which a task has been decomposed, are considered in relation to man-machine interactions. [Swain 83] “decomposes” man into the components of perception, mediation and response; see Figure 5-6. On the one hand, man is *exposed* to external impacts but, on the other hand, he can also *act on* the external environment. For the identification of errors that may occur in these interactions, Swain and Guttman listed a number of factors characterising these five components (external input, perception, mediation, response, human output).

Incorrect human outputs, i.e. the externally observable aspects of human errors, are divided in Swain's handbook into the groups: errors of omission and errors of commission. The sub-categories (Table 5-8) of these main groups are suitable as tools for inductive error identification, for example:

- Category “too much”: What happens if an excessively high shutdown gradient is adjusted?
- Category “too little”: What happens if less than four steam generators are depressurized?
- Category “too early”: What happens if the steam generators are depressurized too early?

Deductive (in fault tree generation, see Figure 5-4) and inductive (checking possible deviations from the correct performance of a task with the error categories listed in Table 5-8) analyses thus ensure a largely complete identification of errors acting on the system as obvious (basically observable) modes of human behaviour. In addition, the man-machine model outlined in Figure 5-6 and the categories listed in Tables 5-9 and 5-11 also help to identify factors contributing to the causation of such errors. Knowledge of these factors plays a role in error quantification, especially in the assignment of human error probabilities (HEPs) and in the identification of performance shaping factors (PSFs) which have an increasing or decreasing effect on these HEPs, see steps 6 and 7 in Table 5-12.

In all cases background knowledge about basic findings in error causation (e.g.: [Swain 83], [Wickens 84], [Reason 90]) and observable circumstances (constellations of external PSFs, context) of errors in real incidents (cf.: [Flanagan 54], [Dougherty 93]) would be very helpful for predictive error identification (see Figure 5-1).

Recent model concepts consider man-machine interactions in a larger context, see Figure 5-5, including also organisational in addition to technical features. It should therefore be examined in a separate research project whether and how these extended models lead to *new* insights into error identification and are quantifiable.

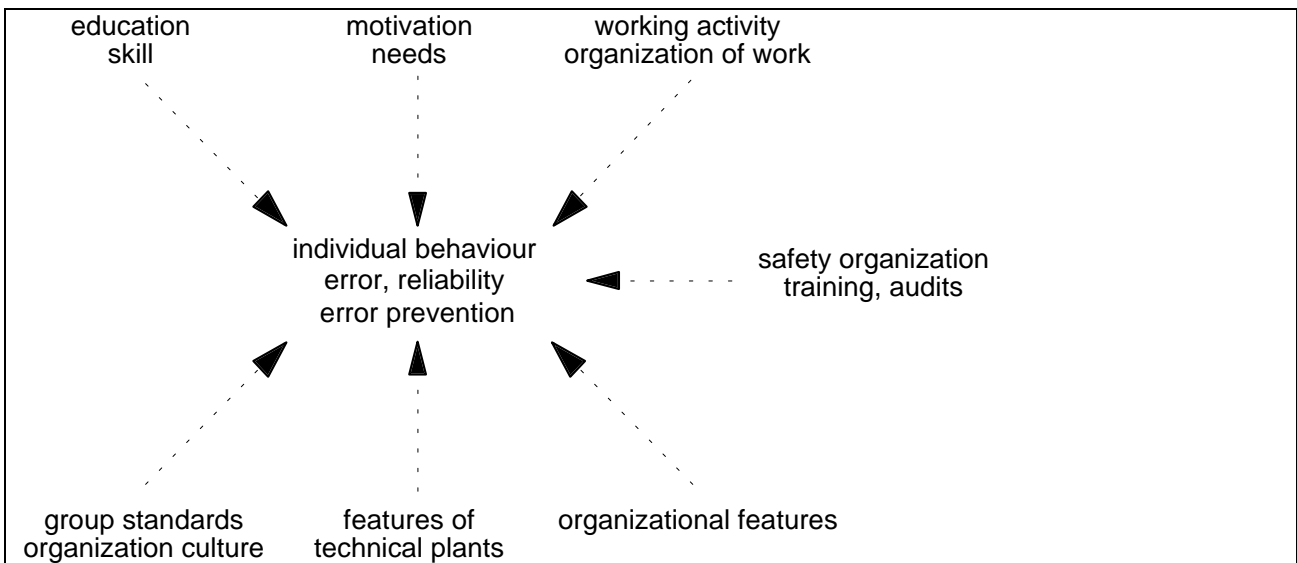


Figure 5-5: Survey of factors influencing individual behaviour (translated from [Grote 93], page 6)

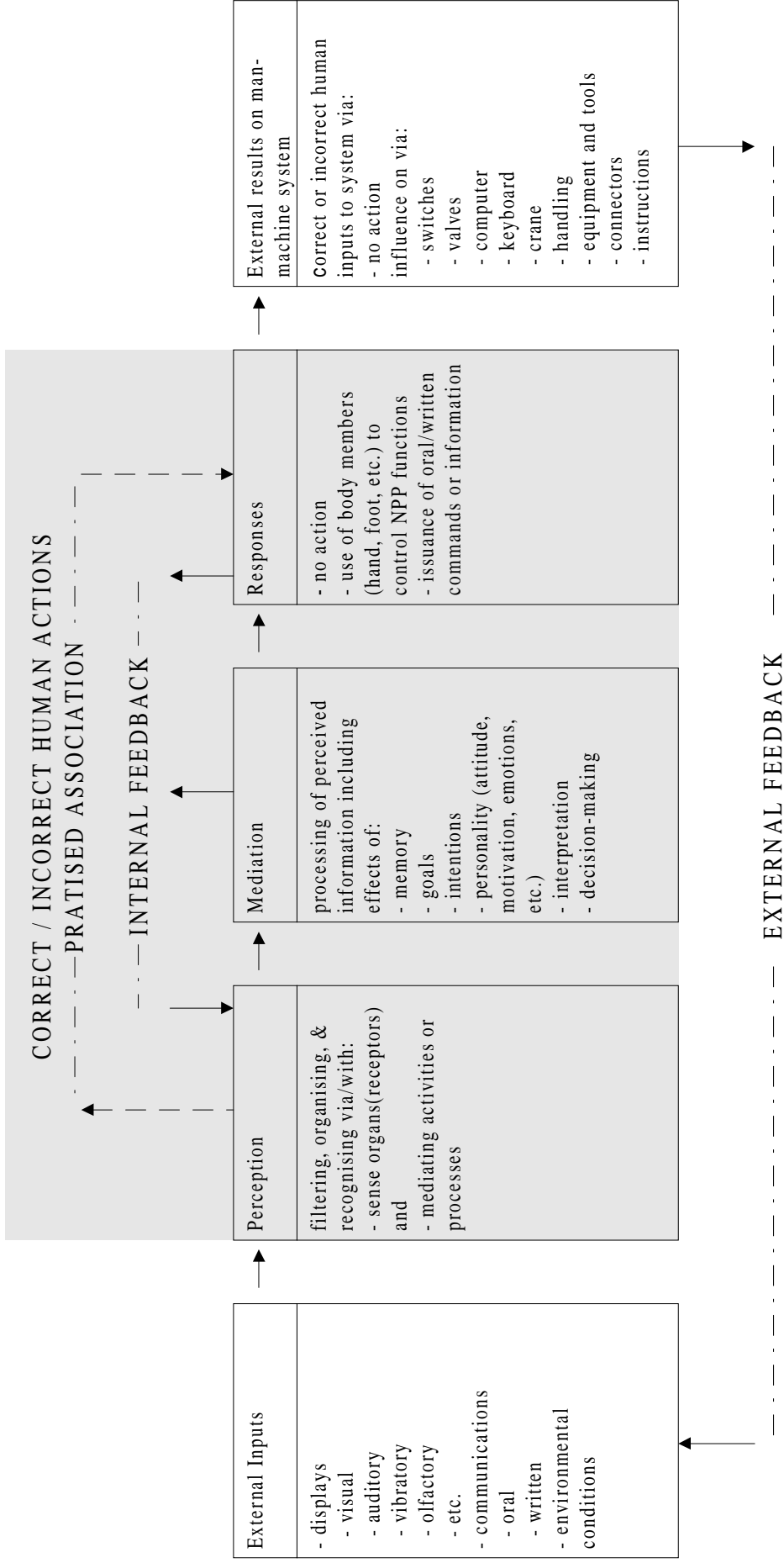


Figure 5-6: Simplified model for the incorporation of human components into a man-machine system From [Swain 83] (page 4-10), slightly simplified

5.4.2 Computer programs as tools

As already mentioned above, the degree of completeness in error identification greatly depends on the analyst's expertise, among other aspects on his experience from previous analyses performed, on his knowledge of plant-engineering details, on his knowledge of cognitive error mechanisms, on his knowledge of previous incorrect actions in technical systems and on his ability to draw analogy conclusions for the system analysed.

Attempts have recently been made to reduce this dependence by developing special computer programs. The computer is to serve as an aid for the analysis of complex man-machine systems. The basic idea of the ESAP computer program (expert system for the systematic analysis of operator actions [Degen 94]) is to provide support in the application of THERP in order to reduce subjective influences, and also to create a knowledge base from previous analyses; this knowledge base can then be used for current analyses. The development of ESAP is not yet completed.

The computer system CES («Cognitive Environment Simulation» [Woods 88], [Woods 90]) is also still at the development stage. Work has progressed to a level where the first test applications are available. CES simulates the process of decision-making; see Figure 5-7.

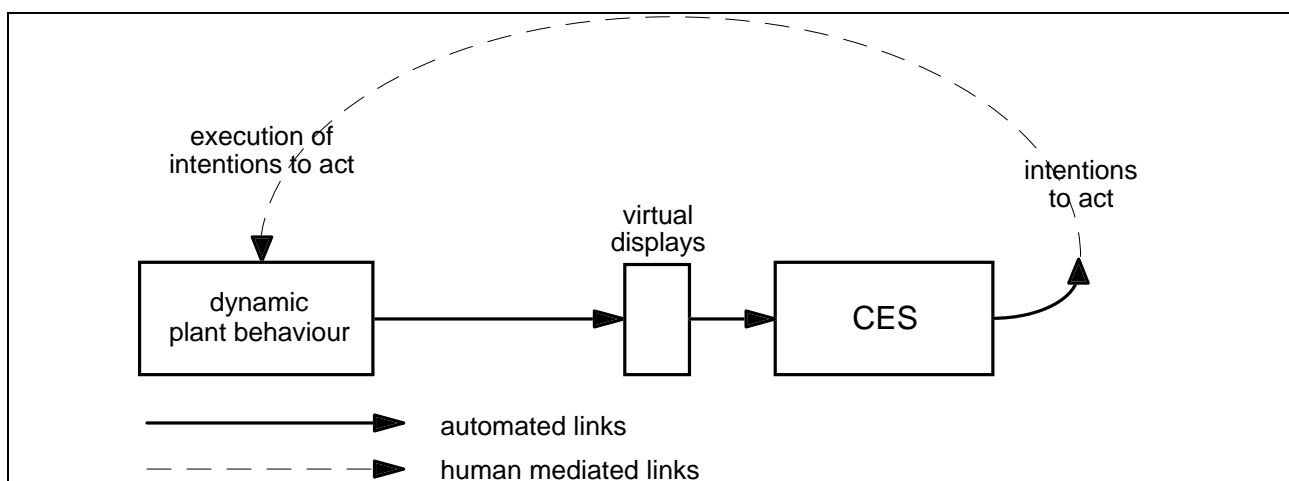


Figure 5-7: Simulation of human intention formation using CES [Woods 88]

Information about dynamic plant behaviour under accident conditions is processed in a manner providing the CES user with a list of intentions to act which could be pursued by the personnel. In simulating operator decisions, insights from psychology and from a number of accidents already occurred are utilized. According to the CES model concept, a decision-based error is caused by a mismatch between problem solution requirements (accident variant, available data, ...) and problem solution capability (knowledge of plant behaviour, preferred strategy in decision-making, ...). Consequently, CES is also in a position to simulate complication factors, i.e. deviations from the text book case of an accident situation. This will be illustrated by an example ([Woods 90], Chapter 4) describing the simulation of decision-making after an accident with steam generator tube rupture:

The CES output showed for the text book case that the personnel diagnoses the accident without any significant delay. In a second case, complications concerning the accident variant were entered *by the CES*

user, for example a leak in the seal of a main coolant pump as an additional failure. The CES output showed an increased diagnosis time. Knowledge deficiency was then entered as a further complication factor, i.e. the personnel attributes the falling pressurizer level caused by the steam generator heating tube leak to the sealing leak. In this case, the CES output showed that accident diagnosis fails.

In this example, the CES user thus decided which complication factors were to be simulated in accident diagnosis. The analyst's expertise is therefore also of decisive significance for the application of such computer programs. According to the present state of the art, these tools can at best support the analyst's knowledge-based skills, but by no means replace them.

5.5 Methods for quantitative human reliability analysis (HRA)

5.5.1 Overview of the variety of HRA methods

This section gives an overview of probabilistic methods for quantitative human reliability analysis. Such methods provide guidelines for the analyst assessing the failure probability of a manual action or the probability of an operator error identified as safety-relevant.

In the meantime, there exists a large variety of quantification methods. PRA-oriented evaluations of individual quantification methods can be found, for example, in [Swain 89] or [Kosmowski 94]. Evaluations from a psychological point of view are to be found in [Reason 90] and [Zimolong 90]. Table 5-14 gives a rough survey of existing quantification methods and some important evaluations concerning these methods.

5.5.2 Selection of significant HRA methods for further evaluation

In addition, two recent methods are also worth mentioning:

- the method based on French operating and simulator experience [Mosneron 90];
- the method derived from the EPRI-sponsored ORE programme (HCR/ORE) [Moieni 94].

A first rough review led to seven methods (Table 5-15) that appear worth evaluating in more detail. This pre-selection is governed by the following three criteria:

1. topicality: new methods for which only few or no evaluations are available (HEART, INTENT, EdF's PHRA, HCR/ORE),
2. PRA relevance: methods frequently used in risk studies (THERP, HCR, SLIM, HEART, EdF's PHRA),
3. cognitive relevance: promising methods for the quantification of nontrivial operator decisions (INTENT, EdF's PHRA, HCR/ORE).

An evaluating description of the standard method THERP (Technique of Human Error Rate Prediction) is given in Section 5.6. The other six methods (EdF's PHRA, HCR, HCR/ORE, SLIM, HEART, INTENT) are briefly described and roughly evaluated in Sections 5.7 and 5.8.

Table 5-14. Methods for human reliability quantification and excerpts from their evaluations.

<i>Method</i>	<i>Evaluation</i>			
	<i>[Swain 89]</i>	<i>[Kosmowski 94] (Chapter 2)</i>	<i>[Reason 90] (Chapter 8)</i>	<i>[Zimolong 90]</i>
THERP Technique for human error rate prediction [Swain 83]	plausible results for competent application; speculative for knowledge-based behaviour during accidents	database available; insufficient consideration of decisions after diagnosis	efficient for competent application; overemphasis of externally observable errors	neglects psychological processes in performing tasks
ASEP Accident sequence evaluation program [Swain 83]	short version of THERP; application inexpensive	enables fast pre-selection of important operator actions	n. e.	n. e.
OAT Operator action tree [Hall 82]	pioneering achievement for diagnosis and time dependence; PSA steps unclear	operator decision analysable	time reliability curves not sufficiently differentiated from a psychological aspect	n. e.
AIPA Accident initiation and progressing analysis [Fleming 75]	only of historical relevance	n. e.	n. e.	n. e.
HCR Human cognitive reliability model [Hannaman 84]	interesting unproven approach; further development necessary	simple handling; incomplete concerning PSFs and time-independent HEPs	time reliability curves not sufficiently differentiated from a psychological view	n. e.
SAINT System analysis of integrated networks of tasks [Kozinski 84], [Wortman 78]	complex interactions can be modelled; models partially difficult to understand; as yet unsuitable for PSA	n. e.	n. e.	n. e.
PC Paired comparison [Seaver 83]	best expert estimation procedure for single HEPs	difficult for HEPs of complex tasks	n. e.	n. e.
DNE Direct numerical estimation [Seaver 83]	requires good reference values; unsuitable for complete HRA	encourages useful discussions among experts	n. e.	n. e.

n. e.: not evaluated

Table 5-14.(cont.) Methods for human reliability quantification and excerpts from their evaluations.

<i>Method</i>	<i>Evaluation</i>			
	<i>[Swain 89]</i>	<i>[Kosmowski94] (Chapter 2)</i>	<i>[Reason 90] (Chapter 8)</i>	<i>[Zimolong 90]</i>
SLIM Success likelihood index methodology [Embrey 84]	flexible; poorly validated; sophisticated	good theoretical background; sophisticated expert estimations	less overemphasis of externally observable errors than in THERP; selection of reference HEPs is critical	PSF interactions are left out of consideration
STAHRA Socio-technical approach to assessing human reliability [Phillips 85]	contributes towards understanding human actions; as yet unsuitable for PSA	not particularly user-friendly at present	n. e.	n. e.
CM Confusion matrix [Potash 81]	pioneering achievement for problems of error in diagnosis	greatly dependent on expert estimations	primarily restricted to qualitative applications	n. e.
MAPPS Maintenance personnel performance simulation model [Siegel 84]	suitable rather for optimisation than for PSA assessment of maintenance work	sophisticated; results difficult to understand	n. e.	n. e.
MSFM Multiple-sequential failure model [Samanta 85]	dependence model, not yet mature	n. e.	n. e.	n. e.
SHARP Systematic human action reliability procedure [SHARP]	useful method for planning HRA in a PSA	useful frame for the integration of several methods	facilitates method selection	n. e.
HEART Human error assessment and reduction technique [Williams 88]	n. e.	simple; excessively isolated consideration of single tasks	n. e.	n. e.
INTENT Method for estimating HEPs for decision-based errors [Gertman 92]	n. e.	promising approach to the quantification of operator decisions	n. e.	n. e.
TESEO Tecnica empirica stima errori operatori model [Bello 80]	n. e.	simple; theoretical background questionable	simple; no hard data base	n. e.

Table 5-15. Significant methods for human reliability quantification.

1.	THERP (“Technique of Human Error Rate Prediction”), described in [Swain 83], short version (ASEP) in [Swain 87]
2.	EdF’s PHRA (Method of Electricité de France for Probabilistic Human Reliability Analysis), described in [Mosneron 89], summarised in [EPS 1300] and [Mosneron 90]
3.	HCR (“Human Cognitive Reliability Model”), described in [Hannaman 84], summarised in [Hannaman 85] and [Hannaman 88]
4.	HCR/ORE (“Human Cognitive Reliability/Operator Reliability Experiments”), based on operator reliability experiments, described in [Spurgin 90] and summarised in [Moieni 94] and [Parry 91]
5.	SLIM (“Success Likelihood Index Methodology”), described in [Embrey 83] and [Embrey 84]
6.	HEART (“Human Error Assessment and Reduction Technique”), described in [Williams 88]
7.	INTENT (“Method for Estimating Human Error Probabilities for Decision-Based Errors”), described in [Gertman 92]

5.5.3 Classification of significant HRA methods

HRA methods may be classified into different classes according to the following criteria:

- the level of detail into holistic and decompositional methods,
- the level of data scale, which is used in the method, into absolute scale, relative scale and ordinal scale,
- the key parameters used by the method into error related, time reliability related and PSF related methods.

5.5.3.1 Classification into decompositional and holistic methods

The classification into decompositional and holistic methods is made according to the procedure, which is used in the assessment process (see [Zimolong 91], [Heslinga 93], [Gerdes 93]). Decompositional methods quantify the reliability of human actions by decomposing a situation into sub-situations up to a defined degree of resolution of the action tree. They assess each single sub-situation, and the reliability for the entire task is inferred by combining the reliability for the sub-tasks. Holistic methods perform a quantitative assessment by assessing the entire situation without distinguishing between the different tasks in a given situation. Table 5-16 allocates the methods to this classification.

Table 5-16. Classification of the considered methods into decompositional and holistic methods

<i>Decompositional Methods</i>	<i>Holistic Methods</i>
THERP ASEP	EdF's PHRA HCR HCR/ORE SLIM HEART INTENT

5.5.3.2 Classification according to the data levels

Another possibility for the classification of HRA methods is the underlying level of the data scale. The level of the data scale may be subdivided into absolute scale, relative scale and ordinal scale.

Absolute scale means that the method does provide real HEP in the range from 0 to 1 where 0 means no error and 1 sure failure. Relative scales are only able to provide information like “if $HEP_1=0.1$ and $HEP_2=0.2$ then event 2 is twice as probable as event 1”. Ordinal scales only can provide information like event 2 is more likely than event 1. It cannot be stated to what extent it is more likely.

A simple rule to distinguish the methods is that absolute scales need real probabilities given by the general probability axiom n/N where n is the observed frequency of events (of interest) and N all observable events. Relative scales need only a subset of the observed events N . The more events observed, the closer the scale is related to the absolute scale. Ordinal scales only need frequencies and no number of observed events. Judgement methods also only provide ordinal scaled data. Normally, methods using this type of scale try to enhance the data level by calibrating the frequencies with some mathematical approach.

Surely, every HRA method intends to achieve the absolute data scale, because it is needed for PSA application. Unfortunately, most of the methods do not provide such data but behave as if they had this data level. A typical example is the SLIM method which uses “anchor tasks” to calibrate other tasks to generate HEPs from data that is of an ordinal scale nature.

Absolute scales are sparse and can be excluded for HRA methods. No HRA method can achieve this data level. Even the THERP catalogue for instance only provides data on the relative scale, because absolute scales need the number of all observable events. Concerning the ordinal scale, it is not usually possible to enhance the data level of methods, because either the underlying assumptions of the mathematical approach or the mathematical approach itself is not proven. Table 5-17 allocates the methods according to the data level. For methods like SHARP for instance, which do not provide data at all, another type of data level is added.

Table 5-17. Classification of the considered methods according to their data scale.

<i>Relative Scale Methods</i>	<i>Ordinal Scale Methods</i>	<i>No Scale at all</i>
THERP	HCR	SHARP
ASEP	SLIM	
EdF's PHRA	HEART	
HCR/ORE	INTENT	

5.5.3.3 Classification according to key parameters

A very practical classification is the distinction of key parameters mainly used by the methods. They may divide the methods into error related, time reliability related and PSF related methods. Table 5-18 gives an overview of the classification.

Table 5-18. Classification of the considered methods according to their key parameters.

<i>Error related</i>	<i>Time reliability related</i>	<i>PSF related</i>
THERP	EdF's PHRA	SLIM
ASEP	HCR	HEART
	HCR/ORE	INTENT

The last classification approach will be used in the detailed description of the following three sections.

5.6 The standard method: THERP

After publication of the preliminary version [Swain 80], the final version of Swain's handbook [Swain 83] on human reliability analysis (HRA) appeared in 1983. These handbooks document a development which began in the early sixties and was referred to as the "Technique of Human Error Rate Prediction" (THERP).

Human reliability analysis using the THERP version described in Swain's handbook is comparatively sophisticated. The American PRA Guide [PRA-PG] specifies approximately two to three man-months per NPP-PRA for the effort required. For this reason, Swain published a short version of THERP [Swain 87] under the title "Accident Sequence Evaluation Program" (ASEP), which permits a time-saving rough assessment ("screening") which, in most cases, leads to more conservative results (higher failure probabilities). Moreover, ASEP also contains explanatory and complementary notes on methodological details in Swain's handbook (1993). [Swain 89] (page 3-41) therefore recommends the use of both sources ([Swain 83], [Swain 87]) for a comprehensive HRA.

The descriptions and evaluations carried out here are based on the 1983 THERP handbook [Swain 83] since THERP and ASEP are based on the same methodological principles. THERP was used in the following studies (cf. [Swain 89], page 3-37):

- Zion PRA [Zion 81];
- Indian Point PRA [Indian Point 82];
- Arkansas Nuclear One PRA [Kolb 82];
- Peach Bottom PRA [Kolaczowski 86];
- Grand Gulf PRA [Drouin 87]
- and in various analyses of non-nuclear systems [Miller 87].

Furthermore, applications are reported in Scandinavian [Hirschberg 90] and German [DRS-B] studies.

5.6.1 *The error model*

On the whole, the THERP error model, which essentially consists of action decomposition, time-reliability correlation and multiplicative PSF model, has a number of attractive features. The decomposition of action provides an important contribution towards understanding the man-machine situation in a complicated technical system. In decomposing an action into meaningful sub-actions perceived by the operator as such, and in thoroughly considering any dependencies involved, important information is obtained about logical and time-dependent correlations of significance for the successful performance of an action.

Moreover, thorough decomposition of an action provides sub-actions which are suitable for experimental verification with respect to their reliability. As a general rule, the determination of experimental reliability is simpler in cases where:

- the error probability (HEP) is higher, i.e. relatively few trials are required in order to obtain a reliable estimate, and
- the number of PSFs to be checked for each trial is small.

The time-reliability correlation as a tool for quantifying the diagnosis is a pragmatic, but meaningful supplement to action decomposition. According to the THERP model concept ([Swain 83], Chapter 12), diagnosis is a highly dynamic (i.e. knowledge-based) activity which is difficult (and with considerable effort) to access by decomposing the action. Figure 5-6 illustrates the dynamic component of a diagnosis activity after the occurrence of an accident. It is to be expected that both the external and the internal control loop (feedback) are 'passed' several times. Simulator experience confirms this model concept:

»Of particular significance from the operators' point of view, and thus their main problem, is the fact that, after the beginning of a major accident, ... a completely unstructured decision problem is initially encountered in the control room. The operator does not know for the moment what is happening....«

(translated from [Hoffmann 89], page 5.2-6).

For the practical feasibility of an HRA it is therefore meaningful to model the diagnosis as *one* action, although it comprises several sub-actions. Whether this is meaningful from a psychological point of view will have to be investigated in more detail. The quantification problems arising in connection with time-reliability correlations will be discussed later.

The multiplicative PSF modification of nominal human error probabilities (NHEPs) is also to be classified as meaningful from a pragmatic rather than psychological point of view. It would be ideal to have an empirically well-founded failure probability under comparable boundary conditions (PSFs) for each sub-action to be quantified in the event tree. A realisation of this ideal case would require a considerable amount of data. An additional data requirement would have to be expected for every new HRA. Modification models are therefore indispensable for the practical feasibility of an HRA. The use of multiplicative models is meaningful in so far as a probability is in the first place a multiplicative variable. The probabilistic implementation of all operations of Boolean algebra can be attributed to multiplications by failure probabilities (P) and success probabilities (Q) in the sense of non-failure probabilities ($Q = 1 - P$).

In strict mathematical terms, the direct application ($HEP = NHEP \cdot PSF$) of multiplicative models for probabilities is only approximately correct, for a probability is not a variable whose multiplicity is defined on a ratio scale; the mathematical principles of the scale theory can be found, for example, in [Coombs 75]. The reason for this is obvious. An HEP cannot assume values greater than 1. For example, $NHEP = 0.5$ and $PSF = 5$ would give $HEP = 2.5$ – a nonsensical result which would have to be interpreted as $HEP = 1$. For $NHEP = 0.5$ and $PSF = 10$ the result would also have to be interpreted as $HEP = 1$, although this case ($PSF = 10$) is to be classified as twice as 'susceptible to failure' as the case with $PSF = 5$ – a violation of the prerequisite for the application of a ratio scale.

It is mathematically correct to perform the multiplicative PSF modification for the quotient, $y = b/a = (1 - NHEP)/NHEP$, of the number b of successes and the number a of failures. This quotient is also referred to as the odds ratio. The multiplicity of this quotient is defined on a ratio scale ($0 \dots \infty$) in the same way as that of its reciprocal value (a/b). The mathematically correct modification of a nominal error probability should therefore take place according to the equation

$$HEP = \frac{1}{1 + \frac{y}{PSF}} = \frac{NHEP \cdot PSF}{NHEP(PSF - 1) + 1}$$

and this equation would also have to be used for the empirical PSF determination:

$$PSF = \frac{y_1}{y_2}$$

where y_1 is the odds ratio in situation 1 and y_2 the odds ratio in situation 2. The PSF expresses the extent to which these two situations differ with respect to their susceptibility to failure.

It is shown in [Reer 93] (p. 128) that the results of the mathematically correct modification procedure, $HEP = 1/(1 + y/PSF)$, differ from the results of the modification procedure usually applied in PRA ($HEP = NHEP \cdot PSF$) in proportion to the closeness of $NHEP$ to 1. On the whole, the usual PSA procedure is acceptable, since it tends to overestimate rather than underestimate the HEP.

In principle, the THERP error model is useful and pragmatic. However, there are some indications that some recovery effects as modeled by THERP may be on the optimistic side. According to psychological findings, as summarised in [Semmer 94], a dependence between diagnosis failure and failure of the respective control mechanisms has to be assumed. This point is discussed in more detail in Section 5.6.7.

5.6.2 *The database*

The validity of the THERP data has been shown to be reasonably satisfactory. However, only few HEPs are directly based on empirical data from nuclear power plants and many HEPs are strictly expert estimates by the authors of the handbook. As a positive factor, these expert estimates are based on sound knowledge concerning man-machine interactions in NPPs. Swain's handbook comprising roughly 700 pages is itself the best support for this assessment. It is still one of the most comprehensive collections of knowledge on human reliability with the emphasis being placed on safety-relevant situations in NPPs.

Moreover, some empirical studies confirm the validity of essential HEPs from THERP within the usual PRA uncertainties. The diagnosis failure probabilities determined by EdF simulator experiments [Villemeur 86] are within the uncertainty bands of the THERP diagnosis model; most of the plots (EdF) are very close to the nominal THERP curve. For critical actions after diagnosis, HEPs determined in simulator studies [Beare 84] are within a factor of 2 of the corresponding THERP HEPs.

However, the validity of the THERP HEPs evaluated here as satisfactory only holds in connection with the associated uncertainty factors ($UF = 2...30$, depending on HEP). A satisfactory validity (with respect to input data quality) of the final result can only be expected if the HEP uncertainties are also considered when using the HEPs.

The uncertainty factors recommended in Swain's handbook in connection with error probabilities are based on conservative conclusions from experiments [Wechsler 52] on the variability of human performances.

The variables used by Wechsler to measure human performance are not failure or success probabilities, but concrete quantitative results from working activities, such as the number of words typed per minute. The extent to which the derived uncertainty factors (in human variability) can be transferred to error probabilities remains difficult to evaluate. Investigations are recommended here based on the methodological principles of mathematical statistics concerning variance.

In any case, Swain & Guttman hold the view that the THERP uncertainty factors cover the variability causes listed in Table A-7 of Appendix A (also [Swain 83], p. 7-9). The above-mentioned simulator studies ([Villemeur 86], [Beare 84]) suggest that this applies at least to the uncertainty types listed under point 1 (lack of NPP-specific data) and point 5 (individual and time-dependent variability). The extent to which the other uncertainty types are covered has not been discussed. Further relevant studies are recommended here. Results from the benchmark exercises [Poucet 88] suggest that Swain & Guttman underestimate the variability causes listed under points 3 (incompleteness) and 4 (analyst).

5.6.3 *The performance shaping factors*

THERP allows for the consideration of numerous plant-specific details. A compilation of such details, with comments, can be found e.g. in [Reer 94d].

The detailed discussion of PSFs of significance for human action in NPPs is one of the principal items of Swain's handbook. Concrete studies as well as psychological and ergonomic findings are quoted for numerous PSFs. Many other methods for HRA implicitly use the discussions and definitions of PSFs that appear in THERP.

Many statements in Swain's handbook are highly topical even today. As early as in 1983, for example, the benefits and disadvantages of symptom-oriented (as an alternative to event-oriented) procedures in accident diagnosis were discussed ([Swain 83], p. 3-82). With regard to symptom-oriented procedures, it is definitely not advisable to readily assume increased diagnosis reliability. The authors recommend thorough reading of the studies quoted in Swain's handbook ([von Herrman 83a], [von Herrman 83b]), in which problems with symptom-oriented procedures are also discussed.

Psychological and cognitive factors of significance in making nontrivial decisions are described in THERP, but the associated quantification problem – especially with respect to knowledge-based performance – is only inadequately solved. The principally known methods of HRA exhibit deficiencies in

this respect (see Section 5.2.3.2). For this reason, the quantification of knowledge-based performance in decision-making will be discussed as a separate evaluation criterion.

5.6.4 Treatment of operator decisions at the knowledge-based level

In principle, THERP allows the quantification of a number of errors related to knowledge-based performances. However, the assumed quantification models, e.g. time-reliability correlation for diagnosis, simple PSF of 2.5 for subsequent actions) are too general and mostly based exclusively on expert estimates. There are no indications for the quantitative implementation of many PSFs related to knowledge-based performance.

[Swain 83] (Chapter 12) admits that the THERP diagnosis model is of a very speculative nature and that THERP exhibits deficiencies in the quantification of operator decisions after diagnosis. In [Swain 88] the event category of diagnosis failure (i.e. no correct diagnosis) is divided into the sub-categories

- no diagnosis and
- wrong diagnosis.

This further development is to be classified as positive. In quantitative application, however, difficulties are to be expected, especially due to a lack of probability data on the occurrence and correction of a wrong diagnosis.

5.6.5 The time reliability correlation (TRC)

Curves are available with which the diagnosis failure can be quantified as a function of the available diagnosis time and the general level of operator knowledge. It is positive to note that the time required for diagnosis is modelled as a random variable, since this is a highly dynamic task (see Figure 5-6). Another positive aspect is that the diagnosis failure probability HEP(t) is modelled as relatively slowly decreasing and with a high uncertainty factor (UF = 30) with increasing time (t). This reflects the extrapolative nature of the model in the range of high t-values; in comparison: the empirical diagnosis times published in [Villemeur 86] are all below 25 min.

A negative feature is that THERP assumes the same diagnosis failure time dependence for all accidents requiring response to a single abnormal event. Simulator experience confirms that accident situations differ with respect to the required scope of diagnosis [Hoffmann 89].

A positive aspect is that THERP allows the quantification of an increased scope of diagnosis due to the occurrence of several abnormal events. The practical application of this option is unclear. Table 5-19 shows, for example, that there may be several criteria for successful diagnosis within an accident sequence. As a general rule, the scope of diagnosis increases in proportion to the severity of the accident. However, Swain's handbook does not contain any clear guidelines for the appraisal of conditions (e.g. failure to start the emergency system) increasing accident severity as additional events in the sense of the 10-minute rule (diagnosis is extended by 10 min per additional event). The study [Woods 82], quoted by [Swain 83] (p. 12-14), from which the 10-min rule was derived, has not been verified.

The inadequate consideration of the time factor for post-diagnosis actions is a further weak point of THERP. There is no clear indication that the time required for unnecessary preceding actions should also be taken into consideration in addition to the time required for the (important) actions necessary to control the accident. A necessary action is mostly the end of a series of activities which also contribute to the total

time required. Furthermore, it is advisable that the time required for the primarily rule-based actions after diagnosis should also be modelled as a random variable. The assumption of a fixed (non-distributed) value t_A in THERP means that the probability for significantly exceeding or remaining below t_A is neglected. This is contradicted by some data which confirm that considerable fluctuation widths with respect to the times required are also to be expected in performing rule-based actions; cf. the time data published in [Haas 82] or [Hannaman 88], e.g. a factor of 20 between the earliest and latest time of performance. The probabilistic modelling of the time dependence of diagnosis and subsequent actions also has the advantage that the success probability of subsequent corrections as a function of the residual time available can be quantified [Reer 1994b] in a more realistic way.

Table 5-19. Hypothetical example concerning the dependence of the scope of diagnosis on accident severity

<i>Event sequence (PWR)</i>	<i>Action to be diagnosed</i>
<ul style="list-style-type: none"> • failure of steam generator feeding • failure of automatic actuation of emergency feedwater system 	<ul style="list-style-type: none"> • manual start of emergency feedwater system
<ul style="list-style-type: none"> • manual start of emergency feedwater system fails 	<ul style="list-style-type: none"> • start of emergency system
<ul style="list-style-type: none"> • start of emergency system fails 	<ul style="list-style-type: none"> • steam generator emergency feeding with mobile fire fighting pump

5.6.6 Handling of dependencies between errors

THERP allows a consideration of dependencies that may occur between operator actions. The dependence model is uncomplicated and its principle methodologically confirmed, since it is based on the beta factor method recognised in the PRA specialist community. The beta factor method is frequently considered to be too pessimistic for high redundancy, but this weak point does not apply to THERP handling. If there is more than one recovery possibility (i.e. for the AND operation of more than two non-success events) the THERP method provides for a repeated consecutive application of the dependence model. This corresponds to the MGL method (“multiple Greek letter”), a method also accepted in the PRA community.

In technical reliability analysis it is a standard procedure to empirically determine the parameters of the dependence model used. This standard is not reached for the THERP dependence model, so that the appraisal is here 'only' satisfactory despite its user-friendly features.

The creation of an empirical database through the dependent occurrence of incorrect actions is therefore an important task for future research in the field of HRA. Due to the flexibility of human behaviour a high research effort must be expected to satisfy PRA needs.

In addition to this, basic investigations should be initiated because the modelling of dependence is important from various aspects of an HRA, e.g. dependence between (among)

- errors (within a given sequence of actions),
- failures of two sequences of actions,

- PSFs,
- times in action performances, or
- required and available time to perform a given task.

5.6.7 *Handling of error correction (recovery)*

THERP recovery factors refer to conditions that may reduce the human error probability initially evaluated for a required action. They are distinct from recovery modeling in the general sense, e.g. actions to recover failed equipment.

A positive feature is the systematic consideration of different recovery possibilities (personnel redundancy, alarm, subsequent procedure step, scanning, walk-around). The recommendations given in THERP for their quantitative application deserve a negative appraisal, especially for recovery actions in accident situations. The dependence on the psychological causes of an error to be recovered is not adequately taken into account. Psychological findings, which can be substantiated by simulator and accident experience, confirm the significance of this dependence type.

Some questions arise with respect to the recovery of diagnosis failures. In the analysis examples [Swain 83] (Chapter 21) the existence of an alarm indicating the manual action to be initiated is quantified as an independent recovery factor for the failure of diagnosis. This is a considerable underestimation of the diagnosis problem in stress situations. On the one hand, this optimistic modelling contradicts in a certain way the finding quoted by [Swain 83] (p. 3-42) himself that operators tend to ignore information under stress. On the other hand, psychological findings indicate a strong dependence between diagnosis failure and failure of the respective control mechanisms. The failure of diagnosis is a highly knowledge-based error. Such failure can be denoted as a mistake because the operators do not implement a correct plan. According to [Reason 90] (p. 62) recovery mechanisms are only effective to a very limited extent in the presence of mistakes. Semmer (1994) gives a very appropriate judgement:

»Once we have made a diagnosis,

- we notice above all information supporting our opinion;
- we frequently ignore information contradicting our opinion;
- we look for explanations fitting our opinion in case we cannot overlook the contradicting information.

...

All these tendencies are enhanced by ... stress« (translated from [Semmer 94]).

These findings can be convincingly substantiated by simulator and accident experience. Not a single error out of 19 in identifying the system status in simulated accidents was recovered by the control room crew themselves [Woods 84]. In the Oyster Creek NPP (1979) the operators did not notice an excessively low water level for 30 min, although this was acoustically signalled by a level alarm [Pew 81]. In the Davis-Besse NPP (1985), too, an acoustic monitor proved ineffective, although it had been specifically added due to the TMI accident. The operators did not notice that the pressurizer relief valve did not close any more after the third opening operation, although closure failure was signalled by the acoustic monitor [NUREG-1154]. Consequently, the recovery effect of an alarm could be quantified with the aid of a dependence model (Section 5.6.6) assuming at least a low level of dependence.

THERP recommends the dependence model for the quantification of personnel redundancy in an accident. However, the personnel recovery possibilities and levels of dependence recommended show a slightly optimistic tendency (see Table A-5). According to the evaluation of the Davis-Besse accident in [Reer 93] (Chapter 2), three of a total of five critical actions were performed without personnel redundancy. In the above-quoted simulator study of [Woods 84] only 50 % of the errors in executing actions after identification of the system status were found to be recovered. Moreover, the assumption that a lower dependence should be assumed for the failure to recover diagnosis errors than for the failure to recover other incorrect actions contradicts the above-quoted psychological findings ([Reason 90], [Semmer 94]); see point 4b in Table A-6).

5.6.8 Handling of uncertainties

The assumed lognormal distribution model is easy to handle and sufficient for PRA purposes. For probabilities, however, it is only mathematically correct to a first approximation.

In principle, a lognormal distribution is suitable for quantifying the uncertainty of a multiplicative variable whose multiplicity is defined on a ratio scale. However, as already mentioned in Section 5.6.1 concerning the multiplicative PSF model, a failure probability (P) or a success probability (Q = 1-P) is only approximately multiplicative.

It would be mathematically correct to model the uncertainty of P via the odds ratio (Y = Q/P), because this is a true multiplicative variable:

$$p = \frac{1}{(1+y)} = \frac{1}{(1+e^x)}$$

where:

Y = lognormally distributed variable with $y = (1-p)/p$ as the median and K as the uncertainty factor;

X = normally distributed variable with $\mu = \ln y$ and $\sigma = (\ln K)/1.645$;

p = point value of the failure probability from Swain's handbook;

K = uncertainty factor of p from Swain's handbook.

This will be illustrated by an example. With $p = 0.1$ as the point value of the failure probability (P), the odds ratio has a point value of $y = 9 : 1 = 9$. The following values are obtained with an uncertainty factor (for Y) of $K = 5$ for the lower bound (LB) and the upper bound (UB) of P:

$$LB(P) = \frac{1}{1+y \cdot K} = \frac{1}{1+9 \cdot 5} \approx 0.022$$

$$UB(P) = \frac{1}{1+\frac{y}{K}} = \frac{1}{1+\frac{9}{5}} \approx 0.36$$

By interpreting the point value of a failure probability from Swain's handbook as the expected value of a beta distribution, with LB as 5 % quantile and UB as 95 % quantile, LB and UB being calculated from the

above equations, the following inconsistencies arising from the treatment of uncertain failure probabilities recommended in THERP can be eliminated:

- Swain & Guttman define the point value (p) of a failure probability as the quotient of the number (a) of errors and the number (n) of error opportunities [Swain 83] (p. 2-17). This point value is interpreted in THERP as the median. According to the probability theory, however, this is an expected value (mean). Since the mean is always greater than the median according to the lognormal distribution assumed in THERP, the mean obtained is systematically too high because the empirical point values are based on the estimate $p = a/n$. Example: for an uncertainty factor of 10, the mean of a lognormal distribution is always higher by a factor of 2.7 than the median.
- Swain & Guttman define the success probability Q (i.e. the probability for the complementary failure event) by the following relation: $Q = 1 - P$ [Swain 83] (p. 2-18). All point calculations carried out with this relation are performed with medians in THERP (see Figure A-1). For lognormally distributed failure probabilities, however, a point calculation, in which the relation $Q = 1 - P$ must be used, only leads to a mathematically correct result if the calculation is carried out with expected values. For a beta-distributed failure probability, this relation can also be used for quantile estimates, since the following symmetry relation is valid:

$$Q_{1-\alpha} = 1 - P_{\alpha}$$

From the above calculated values for the upper bound ($UB = 0.36$) and lower bound ($LB = 0.022$) of the failure probability (with $p = 0.1$ as the point value and $K = 5$ as the uncertainty factor relative to the odds ratio), it is thus possible to directly calculate the values for the upper and lower bounds of the success probability:

$$UB(Q) = 1 - LB(P) = 1 - 0.022 = 0.978$$

$$LB(Q) = 1 - UB(P) = 1 - 0.36 = 0.64$$

This symmetry relation usually also applies to the mean value:

$$q = 1 - p = 1 - 0.1 = 0.9$$

The success probability would thus have an expected value of 0.9 and is within the interval (0.64; 0.978) with 90 % certainty.

In any case, the discrepancies arising from the lognormal distribution do not present a serious weakness of the THERP method because they tend to result in higher error probabilities.

5.6.9 Summary

On the whole, the THERP method is considered reasonable by many practitioners when applied in a competent manner. In order to achieve more realistic quantification of diagnosis, of decisions at the knowledge-based level and time dependencies, it should be examined how elements of more recent methods ([Gertman 92], [Moieni 94], [Reer 93], [Mosneron 90]) can be used for improvements.

5.7 New developments with emphasis on time-reliability correlations in accident diagnosis

The time curves for diagnosis failure probabilities in the THERP model [Swain 83] reflect expert estimates by Swain & Guttman. These expert estimates are based on a consensus achieved at the workshop described in [Oswald 82].

Since then, the data situation in the field of diagnosis reliability has improved. Time curves are available which are based on the simulator studies initiated by EdF (Electricité de France) and EPRI (Electrical Power Research Institute, USA). These studies have led to the development of self-contained methods: EdF [Mosneron 90], HCR [Hannaman 84] and HCR/ORE [Moieni 94], HCR/ORE being a methodological further development of HCR, so that the 1984 HCR model version is now only of historical significance.

These methods will be briefly described and commented on in the following.

5.7.1 EdF's PHRA

The method of EdF for Probabilistic Human Reliability Analysis (PHRA) is essentially based on French simulator studies and on models from THERP (pessimistic diagnosis curves, dependence model) and ASEP (maintenance error model). Note: EdF recommends to use the wording PHRA [Mosneron 94] (instead of only HRA) if HEP quantification is involved.

This section only deals with the quantification of errors in an accident sequence. A relevant detailed description (in French) of the EdF method can be found in [Mosneron 89], brief descriptions and illustrations (in English) in [Mosneron 90], [EPS 900], [EPS 1300] and [Mosneron 94]. The EdF method has been used in [EPS 900] and [EPS 1300].

It is worth noting that the discussion in Sections 5.5 to 5.8 mainly concern the quantitative component of an HRA. Therefore, especially for EdF's PHRA, only a limited view will be presented, since EdF's methodology gives priority to the qualitative component of an HRA:

»Quantification models are not so important for us. We think that knowledge about operator behaviour, habits, attitudes, is the core of PHRA (see [Mosneron 94]). This knowledge is extracted from:

- simulator tests;
- interviews with operators, trainers and operation engineers;
- on site surveys.

So, only a specialist who has acquired this knowledge can perform a good PHRA. Then, quantification models only help him to structure his judgement« [Mosneron 96].

Concerning HEP quantification, the EdF method can be regarded to as a positive new development. However, there is a lack of published information on the procedure to follow in applying the method and in deriving model data (primarily probabilities) from simulator data (observations concerning the behaviour of the personnel).

In comparison to the THERP diagnosis model, the EdF diagnosis model is more advanced since it is based, at least in part, on genuine empirical data. However, the method description does not distinguish clearly between diagnosis and execution time. Some methodological aspects of EdF's PHRA are not totally clear. The case study presented in [EPS 1300] (pages 111-113) demonstrates this issue. Two approaches are presented to quantify the failure to bring the Safety Injection System (SIS) into operation in an accident during shutdown:

1. Given no data about the median time $T_{0,5}$ required to implement the action, the time-dependent diagnosis failure probability is determined by diagnosis curve 2 at $t = 50$ min available; cf. Figure A-2 in Appendix A.
2. Given a data-based estimation of $T_{0,5} = 7$ min, the time-dependent probability of failure to implement the action is determined by the standardised curve 1 at $t/T_{0,5} \cong 7$, also with $t = 50$ min as the input value; cf. Table A-8 in Appendix A.

Although it is not explicitly stated in [EPS 1300], one might conclude from this that the post-diagnosis execution time is considered as negligible in the example. However, the $T_{0,5} = 7$ min estimation (i.e. a 0.5 HEP at 7 min) seems to be unrealistic with respect to diagnosis curve 2 because in this curve the failure probability is 1.0 for times below 10 min. This leads us to the conclusion that the first approach (TRC via diagnosis curve) is more pessimistic than the second approach (TRC via standardised curve).

The EdF execution model itself does not refer to modelling of the possible errors (after diagnosis) in an event tree in the sense of THERP (development of an HRA event tree based on a detailed task analysis). Instead, a simple execution model (consisting of three factors: p_B , K_F , p_{NR}) is presented to apply for every critical action. This approach approximately corresponds to the screening methods outlined in THERP [Swain 83] (Table 20-2) or ASEP [Swain 87] (Table 8-5), see also the comparison between EdF's PHRA and ASEP presented in [Mosneron 93].

Furthermore, EdF's level of action decomposition is more holistic, i.e. less detailed than THERP's. THERP's decomposition of tasks is guided by the model outlined in Figure 5-6 in this report (Figure 4-2 in [Swain 83]). This requires the identification of task-participating man-machine interfaces (displays, controls, etc.) and elements of cognition (e.g. interpretation of a pattern of signals). However, for post-diagnosis execution, the latter is usually neglected or modelled speculatively, see [Swain 83] (page 12-9) and Section 5.6.4 in this report. EdF's PHRA accounts for this variety of PSFs in a simplified manner: choosing between two basic HEPs (p_B), three factors for contextual correction (K_F) and four recovery HEPs (p_{NR}) per critical post-diagnosis action.

In this respect (modelling post-diagnosis execution), THERP may have some advantages over the EdF method, because a detailed event tree is an important tool for the identification of

- functional and time correlations between success and/or non-success events, and
- problems with certain man-machine interfaces or elements of cognition.

However, one should note that:

- execution modelling is closely related to error identification, and for error identification, EdF combines systematic analysis and qualitative simulator data (unfortunately not demonstrated in [EPS 1300], but partially demonstrated in [Mosneron 94]) [Mosneron 95];
- the standard event tree in [EPS 1300] (page 105) may be expanded for certain cases (unfortunately not demonstrated in [EPS 1300]) [Mosneron 96].

The use of qualitative data for error identification is an advanced approach in HRA - especially for the incorporation of unrequired actions.

For screening, qualitative data can be used exclusively to obtain information about the most likely errors. However, it should be noted that screening in error identification inherently produces incomplete results. Not every error-likely situation (e.g. an adverse combination of hardware failures, see [Woods 90] or [Reer 93]) is covered by the base of information (e.g. simulator data) which is used for screening.

5.7.2 *HCR model*

The HCR model ([Hannaman 84]; summarising descriptions in: [Hannaman 85] and [Hannaman 88]) is the result of the second phase of a research project sponsored by EPRI (Electric Power Research Institute) to improve PRA methodology with regard to operator actions.

The first phase comprised the generation of a framework, which was referred to as “Systematic Human Action Reliability Procedure” [SHARP] and published in 1984. [SHARP] recommends an action-type-dependent use of quantification methods (on human reliability); e.g. THERP event tree (“HRA event tree”) for rule-based actions in maintenance activities.

The development of the HCR model is a response to the requirement of methods ascertained in [SHARP] for the quantification of the behaviour of control room personnel in the course of an accident.

The HCR model has so far been used in the following studies [Swain 89]:

- La Salle PRA [Hannaman 86];
- Loviisa PRA [Vuorio 87];
- TMI-1 PRA [Wakefield 87];
- PSA for French 900 MWe reactors [Lanore 87].

Moreover, applications can be found in Scandinavian studies [Hirschberg 90].

The time-reliability correlation is the main error model in the HCR model: the longer the time available for performing a task, the higher is the reliability (success probability) and the lower is the failure probability. Success is defined in the HCR model as the performance of a given task, failure as the “non-response ... similar to non-completion of a predefined task” [Hannaman 88]. The following failure causes are specified as examples:

- “slow in performing”;
- “used another method” (for performing the task);
- “was not alerted”;
- “required additional information”.

It should be remarked that HCR is intended to treat only the diagnosis aspect of an HRA. (This restriction of the application range is not clearly visible from the publications on the HCR model.) In [Hannaman 85] alone several statements can be found (on the application range) which are slightly contradictory:

- Figure 3 on page 347 suggests that the HCR model is conceived for the quantification of diagnosis reliability.
- In contrast, steps 1 and 5 of the survey on page 348 suggest that HCR is to be applied to all situations in which an evaluation of human reliability is required.
- Step 7 of the same survey, however, implies that other models should be used for inadvertent operator errors.
- On the other hand, the task example assessed as HCR on page 350 (manual reactor trip) nevertheless includes manual switching to initiate reactor trip.

Within the scope of the evaluation carried out here it is assumed that the application range conceived for HCR is confined to event A3 of the diagnosis/decision model in Figure A-4; this corresponds to a recent explanation [Moieni 94] (p. 33) concerning the HCR model, where this model is valued as a satisfactory error model in the field of diagnosis.

The HCR frame (Figure A-4) distinguishes between the identification (diagnosis) of the system state required for accident control and selection of the action to be performed. This enables e.g. an explicit incorporation of decision-based errors (despite correct diagnosis) due to conflicts between action alternatives; cf. [Reer 93]. In comparison, the diagnosis in THERP comprises interpretation and decision-making. The HCR model itself, however, does not furnish any probabilities for the decision-based errors B4 and B6 in Figure A-4.

Another basically positive feature is the fact that the event category *No Correct Diagnosis* is divided into the subcategories *Misdiagnosis* (A2) and *No Or Late Diagnosis* (A3). This allows e.g. the incorporation of unrequired operator actions [Reer 93] which have a damage-aggravating effect due to incorrect diagnosis. However, the HCR model does not furnish any probabilities in this case either.

In connection with misdiagnosis, the “frame model” (Figure A-4) of the HCR model is possibly inconsistent. According to the statements in [Hannaman 85] (p. 434) the following applies to the sum of probabilities for events associated with diagnosis (A1: correct diagnosis; A2: misdiagnosis; A3: late or no diagnosis):

$$P(A1) + P(A2) + P(A3) = 1,$$

where P(A3) is determined from the time-reliability correlations (the actual HCR model) in Figure A-3.

Consequently, the event $\overline{A3}$, $P(\overline{A3}) = 1 - P(A3)$, which is complementary to $A3$, would not necessarily be a success event, because it would also comprise the failure event $A2$ (misdiagnosis) in addition to $A1$. This would mean that, in addition to times until correct diagnosis, times until misdiagnosis have also been used without further differentiation to determine the time-reliability correlations. This appears illogical because a correlation favourable for reliability would then result if many misdiagnoses occur within a short time. Moreover, there is a contradiction to the time defined in [Moieni 94] (p. 33) for the determination of the HCR correlations: time between accident occurrence and first correct action. Consequently, the following holds for an error probability derived from the HCR model:

$$P(A3) = 1 - P(A1).$$

The inconsistency shown here is indicative of a basic problem in the use of time data for deriving error probabilities. The concrete causes for the more or less great deviation of a particular time required from the average time required remain hidden from the analyst who does not know the exact boundary conditions of time data acquisition.

Another issue for time-dependent models (such as HCR) is that the time perceived as available by the crew could influence the time of performance in such a way that the crew works particularly slowly or particularly fast. Comparable compensation effects can be corroborated by ergonomic findings [Hacker 86].

Because the HCR model only needs a few parameters, the method gives only a coarse picture of the reality in the plant if no detailed task analysis is performed. In this case, the HCR model does not provide enough face validity ([Fujita 92], [Swain 89], [ACSNI 91]). Moreover, the PSFs are insufficient and modelled coarsely. HCR provides only three different PSFs with a maximum of five subdivisions and the underlying situation cannot be considered adequately (see [Swain 89] and [Fujita 92]).

It should be noted that the normalised time reliability curve, which only shows the relationship of available time and needed time, is only applicable to situations which are covered by the underlying data of the TRC [Sträter 94b]. The absolute magnitude of the time reserve (e.g., hours, minutes or seconds) is not considered.

Furthermore, the time reliability curves for skill-based, rule-based and knowledge-based behaviour are not sufficiently appropriate for the real levels of cognitive behaviour, because the knowledge levels are more a continuum than a dichotomy (see [Wickens 84]). This problem becomes worse due to the fact that the classification into skill-based, rule-based and knowledge-based behaviour completely depends on the experience of the persons in certain knowledge domains and may vary between tasks (see [Bubb 92], page 106). Summarising this point, the HCR model is a strongly simplifying classification of human information processing from the point of view of human engineering and cognitive psychology.

The results from the ORE program [Moieni 94] have resulted in a revision of the method, HCR/ORE. The application of the HCR model version of [Hannaman 84] is therefore no longer recommended by EPRI.

Nevertheless, the HCR model remains a milestone in the quantification of time-dependent error probabilities. The methodological principle is a substantial aid for an analyst facing the problem of estimating the distribution function of a time required for performing a task. In order to comply with the multitude of possible performance situations, it is meaningful to assume a multiparameter distribution function, for example a three-parameter Weibull distribution with the parameters T_0 , η and β . If no time

data are available, the analyst faces the difficult task of estimating the three parameters; especially due to problems with the practical (relative to the performance situation) interpretability of η (scale parameter) and β (form parameter). In the HCR model it is sufficient to classify the type of action (with respect to its cognitive demands on the operator) and to estimate the time requirement to be expected on average. This facilitates the assessment considerably because cognitive demands and average times can be determined on the basis of a task analysis (operator interviews, plant visit, walk-through).

The methodological estimation principle of the HCR model (for the determination of distribution functions of action times) is therefore an essential contribution towards reducing the data problem in human reliability analysis. In [Reer 94b] the HCR principle was used to extend THERP so as to also model the times of actions after diagnosis as probabilistically distributed variables.

5.7.3 *HCR/ORE method*

The following brief description of the HCR/ORE method is based on the presentations in [Moieni 94] and [Parry 91]. Further details on the method and data base can be found in [Spurgin 90] and [Parry 92]. The data themselves are the property of EPRI and not readily available to the public; cf. [Moieni 94] (p. 39).

The method described in [Moieni 94] is often designated HCR/ORE because the base methodology includes generic data collected in the operator reliability experiments (ORE) and reflects the lessons obtained from these [Hannaman 84].

The HCR/ORE method can be used for the probabilistic quantification of actions in an accident. It is possible to quantify actions with predefined procedure (“interactions that are dictated by procedures”) as well as actions without predefined procedure (“scenario-specific recovery of failed equipment or realignment of systems”) ([Moieni 94], p. 35); see also Section 5.2.2.3 in this report.

The division of diagnosis failure into the events (1) failure to formulate response and (2) response too late (Figure A-5) is, in principle, to be regarded as an advance in relation to THERP where both events are considered in one model. A separate consideration of the first event (failure to formulate response), however, is required for reasons of plausibility alone, because the lognormal distribution model (HCR/ORE) would lead to unrealistically small error probabilities for large values of available time.

It should be noted that the two events cannot always be clearly differentiated. The error mechanisms (e.g. misreading) specified in [Moieni 94] for the first event (failure to formulate correct response) could, in principle, also contribute only towards a delay in diagnosis without causing complete failure. Such error mechanisms would then be covered, in principle, by the time-reliability correlations and would not have to be specifically modelled. In a strict sense, only those error mechanisms should be considered for quantifying the first event (failure to formulate correct response in Figure A-5) which are irreversible, i.e. ultimately lead to failure of the requested action.

For the quantification of the second event (response too late in Figure A-5) the HCR/ORE method recommends the use of time-reliability correlations defined on a normalised time scale ($t/t_{0,5}$). It is pointed out in [Sträter 94b] (page 8) that the absolute time reserve is not taken into account. This criticism is particularly true if a task is investigated whose average time requirement ($t_{0,5}$) is of an order of magnitude not covered by the HCR/ORE data base. Let us assume that the $t_{0,5}$ values of the HCR/ORE data are times in the minute range. Further investigations are then required in order to clarify whether and how tasks with a time requirement in the range of seconds or hours can be quantified with these data.

Practical application of the method for quantifying actions without predefined procedure (recovery actions) requires the clarification of further methodological details, in particular with respect to the quantification of:

- dependencies between the failure of an action *with* predefined procedure and the failure of an action *without* predefined procedure and
- conflicting decisions if the crew has to choose between an action with and without predefined procedure.

HCR/ORE's simplified method ($p_{NR} = p_I + p_E - p_I p_E$) for quantifying recovery actions provides a feasible screening approach to emergencies during shutdown states of NPPs. The respected criterion (not dictated by procedures) seems to be typical of actions which are required to control emergency situations during plant shutdown, because:

- the identification (by the crew) of a correct action depends on a variety of conditions, some of which are difficult to predict in HRA, e.g. the set of systems taken out of operation for maintenance until emergency, and
- there is often a variety of correct actions which could be chosen by the crew.

HCR/ORE offers a coarse approach to deal with such inponderabilities in HRA.

A separate outgrowth of the ORE experiments is the “decision tree” method, which represents a complementary approach to evaluating HEPs based on combining expert judgement and data concerning the causal factors of deviations observed in simulators. This method is briefly described in Chapter 8, Section 8.3.2.

5.7.4 Summary (time-reliability correlations)

The TRC approach is usually applied for analysing the diagnosis of an abnormal event. In this context, a TRC-derived HEP is the probability $pr(T > t)$ that the required time (T) for diagnosis is greater than the available time (t) for diagnosis. Table 5-20 summarises some TRCs of the methods THERP, EdF's PHRA, HCR and HCR/ORE.

Table 5-20. Comparative evaluation of HRA methods according to percentiles of the time required for the diagnosis of an abnormal event. The HCR/ORE percentiles are based on imprecise readings from ([Moieni 94], Figure 6), the response types are defined in Table A-9.

HEP	0.5	0.1	0.01	0.001	0.0001	0.00001
THERP, nominal model	2 min = $T_{0.5}$	10 min = 5 $T_{0.5}$	20 min = 10 $T_{0.5}$	30 min = 15 $T_{0.5}$	60 min = 30 $T_{0.5}$	1500 min = 750 $T_{0.5}$
THERP, initial-screening model	10 min = $T_{0.5}$	20 min = 2 $T_{0.5}$	30 min = 3 $T_{0.5}$	60 min = 6 $T_{0.5}$	1500 min = 150 $T_{0.5}$	≈ 3000 min (extrapolation, not tabled in [Swain 83])
EdF, curve 1	≈ 3 min = $T_{0.5}$	≈ 9 min = 3 $T_{0.5}$	≈ 21 min = 7 $T_{0.5}$	For $t > 30$ min both curves remain at a constant residual HEP of 0.005		
EdF, curve 1'	≈ 9 min = $T_{0.5}$	≈ 12 min = 1.33 $T_{0.5}$	≈ 21 min = 2.33 $T_{0.5}$			
EdF, curve 2	12 min = $T_{0.5}$	20 min = 1.67 $T_{0.5}$	30 min = 2.5 $T_{0.5}$	For $t > 50$ min (approximately) the curve remains at a constant residual HEP of 0.005		
EdF, curve 3	22 min = $T_{0.5}$	30 min = 1.36 $T_{0.5}$	For $t > 300$ min (approximately) the curve remains at a constant residual HEP of 0.03			
HCR, knowledge	$T_{0.5}$	2.74 $T_{0.5}$	5.84 $T_{0.5}$	9.34 $T_{0.5}$	13.19 $T_{0.5}$	17.27 $T_{0.5}$
HCR, rule	$T_{0.5}$	2.12 $T_{0.5}$	3.88 $T_{0.5}$	5.75 $T_{0.5}$	7.68 $T_{0.5}$	9.68 $T_{0.5}$
HCR, skill	$T_{0.5}$	1.52 $T_{0.5}$	2.15 $T_{0.5}$	2.74 $T_{0.5}$	3.29 $T_{0.5}$	3.82 $T_{0.5}$
HCR/ORE, response type 1	$T_{0.5}$	PWR: ≈ 2 $T_{0.5}$ BWR: ≈ 2.55 $T_{0.5}$	PWR: ≈ 3.6 $T_{0.5}$ BWR: ≈ 5.5 $T_{0.5}$	PWR: ≈ 5.5 $T_{0.5}$ BWR: ≈ 9.6 $T_{0.5}$	PWR: ≈ 8.1 $T_{0.5}$ BWR: ≈ 16 $T_{0.5}$	PWR: ≈ 9 $T_{0.5}$ BWR: ≈ 18.5 $T_{0.5}$
HCR/ORE, response type 2	$T_{0.5}$	PWR: ≈ 1.6 $T_{0.5}$ BWR: ≈ 2.2 $T_{0.5}$	PWR: ≈ 2.3 $T_{0.5}$ BWR: ≈ 4.1 $T_{0.5}$	PWR: ≈ 3.05 $T_{0.5}$ BWR: ≈ 6.6 $T_{0.5}$	PWR: ≈ 3.9 $T_{0.5}$ BWR: ≈ 10.2 $T_{0.5}$	PWR: ≈ 4.2 $T_{0.5}$ BWR: ≈ 11.5 $T_{0.5}$
HCR/ORE, response type 3	$T_{0.5}$	PWR: ≈ 2.65 $T_{0.5}$ BWR: ≈ 2.75 $T_{0.5}$	PWR: ≈ 5.9 $T_{0.5}$ BWR: ≈ 6.3 $T_{0.5}$	PWR: ≈ 10.55 $T_{0.5}$ BWR: ≈ 11.6 $T_{0.5}$	PWR: ≈ 18 $T_{0.5}$ BWR: ≈ 20.1 $T_{0.5}$	PWR: ≈ 20.9 $T_{0.5}$ BWR: ≈ 23.6 $T_{0.5}$

The percentiles (listed in Table 5-20) must be seen in the light of additional features of the methods. One feature concerns the limitation of the TRC approach for the extreme case that the available time (t) is very large. EdF's PHRA covers such a case by a "residual" HEP which remains constant beyond a certain time limit. HCR considers "faulty detection or cognitive processing" as an additional (not covered by TRC), time-independent diagnosis error. In HCR/ORE such time-independent error is denoted as "failure to formulate correct response" (see Figure A-5), and has to be quantified separately. THERP gives no explicit time-independent or residual HEP for diagnosis. However, de facto, the HEP of 10^{-4} at $t = 60$ min (THERP, nominal model) could be interpreted as residual, because beyond 60 min the HEP decreases at a negligible rate of one order of magnitude per day. Nevertheless, compared to $5 \cdot 10^{-3}$ (EdF, curve 1 or 1'), 10^{-4} is a very optimistic residual HEP.

Furthermore, Table 5-20 shows that, in contrast to TRC models for an absolute time scale (THERP, EdF's PHRA), normalised time scale models need a direct numerical estimation of the median diagnosis time ($T_{0,5}$). On the one hand, this is an advantage because such an option for estimation results in high flexibility. However, on the other hand, such an option produces two elements of uncertainty:

1. the estimation itself, especially if no timing data are available;
2. inadequate consideration of the absolute time reserve ($t - T_{0,5}$).

In HRA practice, another problem of HCR is that the behaviour levels (knowledge, rule, skill) do not sufficiently correspond to real levels of cognitive behaviour (see Section 5.2.3.2). However, due to this point, the ORE program resulted in HCR/ORE - a modified version of HCR.

New developments with emphasis on expert estimates structured according to performance shaping factors

The lack of situation-specific data on human reliability leads to a situation where the analyst frequently has to estimate a particular error probability (p). There are a number of methods for estimating p not directly, but indirectly via performance shaping factors (PSFs). The advantage of these methods is an increased transparency of a subjective estimate. An essential disadvantage is the additional uncertainty of the mathematical model used for the PSF-dependent calculation of p . Even 'demathematized' models (such as the decision tree in Figure A-6) exhibit this uncertainty since the mathematical operations have already been performed in advance. The different mathematical principles for the PSF-dependent derivation of an error probability p will be briefly discussed here using the example of three methods: SLIM [Embrey 84], HEART [Williams 88] and INTENT [Gertmann 92].

5.8.1 *SLIM*

This section discusses the SLIM method published in its preliminary version in [Embrey 83] and in its final version in [Embrey 84]. It should be noted that a large number of "full-scope" HRA applications, particularly in the U.S., use not Embrey's SLIM, but instead a variant developed by PLG, Inc. This variant, referred to here as **PLG SLIM**, addresses some of the criticisms presented here and is discussed in Section 5.8.2.

One problem occurs if the SLIM model contains a PSF of decisive effect. The available time is such a PSF. Let us assume that there is not enough time available for practically completing the task in time; a THERP quantification would lead to a failure probability of 1.0 in this case. This logically correct consistency is not given for the SLIM model because the maximum effect of the PSF 'time' is limited by its weighting factor. The calculations performed in Table 5-21 illustrate this weakness of SLIM.

Table 5.21. Example calculation for the sensitivity of a SLIM result when changing the available time

Task [Embrey 84] (Vol. I, Section 3.1)	Diagnosis of a transient and correct response to failure of an emergency feedwater pump	
PSFs	1) quality of information available in the control room 2) quality of procedures 3) time available for diagnosis and response 4) training level	
weighting factors (for the 4 PSFs)	$\underline{w} = (0.50; 0.25; 0.15; 0.10)$	
PSF ratings: 0 = as unfavourable as credibly possible 1 = as favourable as credibly possible	$\underline{x} = (0.7, 0.2, 0.1, 0.5)$ [Embrey 84], Vol. I, p. 6)	$\underline{x} = (0.7; 0.2; 0; 0.5)$ (changed rating for PSF 3)
success likelihood index	$SLI = 0.5 \cdot 0.7 + 0.25 \cdot 0.2 + 0.15 \cdot 0.1 + 0.1 \cdot 0.5 = 0.465$	$SLI = 0.5 \cdot 0.7 + 0.25 \cdot 0.2 + 0.15 \cdot 0 + 0.1 \cdot 0.5 = 0.450$
model (decimal logarithm)	$\log(1 - p) = 0.00655 \cdot SLI - 0.00567$	
failure probability	$p = 0.0060$	$p = 0.0063$

Table 5-22. The linear dependence of the logarithmic success probability (log (q)) is not equivalent to the linear dependence of the logarithmic error probability (log (p)); (Based on [Kosmowski 94b])

Task [Embrey 84] (Vol. I, Section 3.1)	Diagnosis of a transient and correct response to failure of an emergency feedwater pump	
success likelihood index	SLI = 0.465	
reference error probabilities with associated SLI values	$p_1 = 0.001, SLI_1 = 0.8$ $p_2 = 0.01, SLI_2 = 0.2$	
model variant	[Embrey 84] (Vol. I, p. 6) (q-variant of SLIM) $\log(q) = \log(1 - p)$ $= a \cdot SLI + b$	[Embrey 83] (p. 15) (p-variant of SLIM) $\log(p) = a \cdot SLI + b$
model parameters	$a = \frac{\log\left(\frac{1 - p_1}{1 - p_2}\right)}{SLI_1 - SLI_2}$ $= 0.00655$ $b = \log(1 - p_1) - a \cdot SLI_1$ $= -0.00567$	$a = \frac{\log\left(\frac{p_1}{p_2}\right)}{SLI_1 - SLI_2}$ $= -1.67$ $b = \log(p_1) - a \cdot SLI_1$ $= -1.66$
failure probability of the task (with SLI=0.465)	$p = 1 - 10^{a \cdot SLI + b}$ $= 0.0060$	$p = 10^{a \cdot SLI + b}$ $= 0.0037$

Moreover, a special feature to be criticised in SLIM is the fact that the mathematical model is not uniformly represented: sometimes the logarithmized success probability (q-variant) is the target parameter and sometimes the logarithmized error probability (p-variant). Table 5-22 shows that these two model variants are not equivalent to each other. For the same initial assumptions, the analysis example with the q-variant gives an error probability higher by a factor of 1.6 than with the p-variant. The methodology descriptions available for SLIM are thus inconsistent.

This contradiction (Table 5-22) has a deeper cause which was already dealt with in Sections 5.6.1 and 5.6.8. It was shown in [Reer 93] (Section 5.5.2.3) that the p-variant of the SLIM model is equivalent to a multiplicative model

$$p = p_0 \cdot PSF_1 \cdot PSF_2 \cdots$$

in which the wanted error probability p can be represented as the product of a nominal value (p_0) and one factor or several factors. As already discussed in Section 5.6.1, such a model is only approximately mathematically correct because a probability in the strict sense is not a multiplicative quantity. It is mathematically correct if the odds ratio, $y = (1-p)/p$, or the error chance ratio ($1/y$) is modelled as a multiplicative target quantity. Table 5-23 shows that a consistent model is thus obtained. The same result is obtained with the logarithmized odds ratio as with the logarithmized error chance ratio.

Table 5-23. The linear dependence of the logarithmized error chance ratio is equivalent to the linear dependence of the logarithmized odds ratio

Task [Embrey 84] (Vol. I, Section 3.1)	Diagnosis of a transient and correct response to failure of an emergency feedwater pump	
success likelihood index	SLI = 0.465	
reference error probabilities with associated SLI values	$p_1 = 0.001, SLI_1 = 0.8$ $p_2 = 0.01, SLI_2 = 0.2$	
reference odds ratios	$y_1 = (1 - p_1) / p_1 = 999$ $y_2 = (1 - p_2) / p_2 = 99$	
model variant	$\log\left(\frac{1}{y}\right) = a \cdot SLI + b$	$\log(y) = a \cdot SLI + b$
model parameters	$a = \frac{\log\left(\frac{1/y_1}{1/y_2}\right)}{SLI_1 - SLI_2}$ $= -1.67$ $b = \log\left(\frac{1}{y_1}\right) - a \cdot SLI_1$ $= -1.66$	$a = \frac{\log\left(\frac{y_1}{y_2}\right)}{SLI_1 - SLI_2}$ $= 1.67$ $b = \log(y_1) - a \cdot SLI_1$ $= 1.66$
odds ratio and failure probability of the task (with SLI=0.465)	$y = 1 / 10^{a \cdot SLI + b}$ $= 273$ $p = 1 / (1 + y)$ $= 0.0037$	$y = 10^{a \cdot SLI + b}$ $= 273$ $p = 1 / (1 + y)$ $= 0.0037$

Like every other method that is using expert judgement, SLIM is exposed to additional uncertainties in the data due to the subjective judgement of the expert. Therefore, the danger of sparse consistency in the results is high. This effect may be intensified by the fact that the qualification of the experts performing the judgement is not defined (see [Berg 92], [Swain 89], [Reichart 85]).

An additional problem arises from the calibration procedure, which only needs two anchor tasks. Firstly, the theoretical basis of the calibration is not validated ([Swain 89], [Bubb 92]). Secondly, uncertainties in the parameter estimation of the SLI might be high due to the fact that no aids are given for the determination of the error probabilities of the anchor tasks or for the determination of the similarity of the two tasks which are used as anchors (see [Swain 89]). If only two tasks are taken, the variation of the calibration data cannot be calculated.

5.8.2 “PLG SLIM”

The version of SLIM developed by PLG, Inc., is based on Embrey’s SLIM [Embrey 84] and differs from SLIM primarily in terms of implementation. The main reference for the implementation discussed here is [Chien 88]¹; recently, a more complete description of the method appeared in the Surry Low Power and Shutdown Study [Musicki 94]. Unfortunately, most evaluation/validation studies refer to Embrey’s version [e.g. Poucet 89, Zimolong 91].

Some of the key differences between *PLG SLIM* and Embrey’s *SLIM (original SLIM)* include:

- The likelihood index is formulated in failure space (FLI) rather than success space (SLI). As a result, PLG SLIM versions are sometimes referred to as **FLIM** methods.
- Judgements are elicited from experts familiar with the tasks, typically operators and training personnel. SLIM and SLIM-MAUD are generally intended for the use of a single HRA expert (analyst) or group of interacting experts. The calibration process in PLG SLIM is performed by the HRA analyst.
- A set of PSFs appropriate for nuclear power plant PSA applications are suggested as part of the implementation. The task experts assign an importance to these pre-selected PSFs in the form of weights.
- Because of the involvement of multiple experts in the PLG SLIM application, and to enhance the consistency of application, scaling guidance for the PSF ratings is provided. The optimal points of the PSF rating scales are systematically defined to be at the rating of 0 on the 0-10 scale; this avoids the controversial “re-scaling about an ideal point” required in the original SLIM.

The SLIM family of methods has the advantage of avoiding the direct estimation of probabilities by the task experts, a process generally recognised to be unreliable due to judgement biases. SLIM methods provides an approach for a systematic and structured elicitation of PSF ratings, a task better suited for expert judgement.

The critical aspect of the method is the calibration process. The index (FLI) based on the PSF judgements yields only a relative likelihood (a ranking) of the task HEPs. The calibration process is required to

1. It should be noted that some SLIM applications performed by PLG, Inc., differ from the implementation described in [Chien 88]. In particular, the U.S. studies for the Diablo Canyon PRA (19) and the Oyster Creek PSA both used a method with 21 PSFs rather than the 7 typical of PLG SLIM.

convert the relative scale to HEPs; it requires either tasks with “known” HEPs or “accepted” HEPs.² The general scarcity of validated HEPs makes calibration a problem.³

However, it may be possible to derive plant-specific HEPs for calibration using other methods (considering their weaknesses) [Dang 95]. In such an approach, the use of these methods to derive a small number of calibration values would be supported by more detailed analyses; such analyses might be too resource-intensive to apply for all actions considered in a PSA.

5.8.3 HEART

Like SLIM, the HEART method [Williams 88] is an approach based on the selection of PSFs of significance for the success or failure of an operator task investigated. It was utilised at the task-based level (e.g. 'fail to achieve a certain task'), rather than at the elemental level (e.g. 'fail to close valve') as THERP may be utilised [Kirwan 95]. The HEART method was used in various British studies [Kirwan 94].

For detailed HRA, the use of HEART is subject to some limitations. However, elements of HEART (e.g. incorporation of certain EPCs) can be used for sub-tasks or errors identified in the course of an action decomposition and for which no situation-specific error probabilities are available; compare also the evaluation of SLIM in Section 5.8.1 of this report. For the quantification of holistic tasks (in the sense of system functions), HEART should be considered as a screening method.

HEART provides the analyst with data on nominal error probabilities (NHEPs) and performance shaping factors (PSFs), called 'error producing conditions' (EPCs). The descriptions of the NHEPs are rather general. In addition, quantitative values are provided for the influence of the error producing conditions (EPCs).

It is thus assumed that an EPC selected as being relevant is not yet covered by the selected NHEP. A potential problem is that the data base from which the NHEP was determined may contain cases where this EPC was involved in the causation of errors. This would then lead to double consideration. More details are given in Table 5-24.

Table 5-24 illustrates the uncertainties which arise from the NEHP selection alone. The input values are identical. Nevertheless, HEART produces an HEP result which is higher by a factor of 1.35 than the SLIM result. In [Reer 93] (Table 45) it is shown that

- HEART is always more pessimistic than the p-variant of SLIM , and
- the differences (between the HEP results) are negligible for small values of K.

However, further studies are required in order to clarify which model is more correct according to psychological theory.

-
2. This comment naturally applies to any method where relative likelihoods are determined, for instance, for the Paired Comparisons method (not discussed here).
 3. The calibration required in applying the SLIM family of methods is problematic for studies comparing HRA methods, such as the Human Factors - Reliability Benchmark Exercise (HF-RBE) [Poucet 89, Zimolong 91]. Values obtained with SLIM are completely dependent on the calibration data.

Table 5-24. Quantitative comparison of HEART and SLIM

Error [Williams 88]	Failure to isolate a bypass route	
nominal human error probability	$p_0 = 0.003$	
error producing conditions (EPCs)	<ol style="list-style-type: none"> 1. inexperience 2. opposite technique 3. risk misperception 4. conflict of objectives 5. low morale 	
total (maximum) multiplicative effects of the EPCs: $K_1 \dots K_5$	$\underline{K} = (3; 6; 4; 2.5; 1.2)$	
ratings (assessed proportions of the EPC effects): $x_1 \dots x_5$	$\underline{x} = (0.4; 1.0; 0.8; 0.8; 0.6)$	
model	$p = p_0 \prod_{i=1}^5 [1 + (K_i - 1) \cdot x_i]$ (HEART)	$\log(p) = a \cdot SLI + b$ (SLIM, p-variant)
assessed HEP	$p = 0.003 \cdot 1.8 \cdot 6.0 \cdot 3.4 \cdot 2.2 \cdot 1.12 = 0.27$	$p = 10^{a \cdot SLI + b} = 0.2$ with# $SLI = \sum_{i=1}^5 [w_i (1 - x_i)] = 0.22$ $w_i = \log(K_i) / \sum_{i=1}^5 \log(K_i)$ $a = - \sum_{i=1}^5 \log(K_i) = -2.33$ $b = \log(p_0) - a = -0.19$

See [Reer 93] (Section 5.2.3.3) or [Reer 94c] for mathematical details.

5.8.4 INTENT

INTENT is a topical new development in the field of quantification methods based on PSF-structured expert estimates. The publication discussed here [Gertman 92] does not refer to any PSA application of INTENT. INTENT is specialised in the quantification of decision-based errors due to errors of intention. INTENT is a promising approach complementing the existing methods with respect to the quantification of non-trivial decisions by the personnel.

INTENT is not intended to be a complete quantification method but is instead conceived for filling gaps which other quantification methods exhibit in the field of decision-based errors.

The version of INTENT described in [Gertman 92], however, is difficult to use in practical PSA applications. The definitions concerning nominal errors and performance shaping factors are lacking in clarity. Furthermore, there are no explanations concerning the error mechanisms relating to the PSFs, for example »safety culture«.

Table 5-25 shows that the mathematical model underlying the INTENT method leads to approximately the same result as the mathematical model of the p-variant of the SLIM method. On the contrary, the result from the q-variant of SLIM is higher by about a factor of 3 – see also Table 5-22.

Table 5-25. Quantitative comparison between INTENT and SLIM

Error (example in [Gertman 92])	Non-use of the formulas (provided inside the procedures) for technical calculations		
success likelihood index	SLI = 0.39		
calibration points (reference HEPs and respective SLI values)	$p_1 = LB = 0.0016, SLI_1 = 0.95$ $p_2 = UB = 0.047, SLI_2 = 0.05$		
model	(INTENT) $\ln(p) =$ $\sigma \cdot \phi^{-1}(1 - SLI) + \mu$	(SLIM, q-variant) $\ln(q) = \ln(1 - p)$ $= a \cdot SLI + b$	(SLIM, p-variant) $\ln(p) =$ $a \cdot SLI + b$
model parameters	$\sigma = \frac{\ln(UB / LB)}{3.29}$ $= 1.03$ $\mu = \frac{\ln(UB \cdot LB)}{2}$ $= -4.75$	$a = \frac{\ln\left(\frac{1 - p_1}{1 - p_2}\right)}{SLI_1 - SLI_2}$ $= 0.052$ $b = \ln(1 - p_1)$ $- a \cdot SLI_1$ $= -0.051$	$a = \frac{\ln\left(\frac{p_1}{p_2}\right)}{SLI_1 - SLI_2}$ $= -3.76$ $b = \ln(p_1) - a \cdot SLI_1$ $= -2.87$
HEP (for SLI=0.39)	$p = e^{\sigma \cdot \phi^{-1}(1 - SLI) + \mu}$ $= 0.011$	$p = 1 - e^{a \cdot SLI + b}$ $= 0.03$	$p = e^{a \cdot SLI + b}$ $= 0.013$

The attempt made in INTENT to classify decision-based errors according to their psychological causes rather than according to the external accompanying circumstances is a positive development.

5.8.5 Summary

PSF-oriented quantifications require

- judgements (i.e. ratings, in some methods together with weightings) according to performance shaping factors (PSFs, wording used in SLIM and INTENT) or error producing conditions (EPCs, wording used in HEART), and
- calibration of the corresponding mathematical model.

According to the items above, SLIM shows a high flexibility on the one hand and requires sophisticated work by the user of the method on the other hand.

The SLIM user is free in selecting or assessing:

- the relevant PSFs,
- their ratings,
- their weights, and
- at least two reference HEPs for calibration.

Compared to a user of SLIM, a user of HEART or INTENT has reduced degrees of freedom on the one hand and less sophisticated work to do on the other hand.

In HEART,

- a subset of EPCs has to be selected from a predefined list of 38 EPCs,
- ratings (according to proportions of effect) are required for the selected subset of EPCs, and
- calibration is determined by the assignment of one suitable NHEP (out of a list of nine NHEPs presented for generic described tasks).

INTENT needs

- ratings for 11 preselected (and pre-weighted) PSFs, and for calibration,
- the selection of one error category (out of a list of 20 predefined categories).

As with other HRA methods, the SLIM, HEART and INTENT methods neglect the potentially important dependencies between PSFs. A model simple to handle for the provisional solution of this problem is proposed in [Reer 93] (Section 5.2.3, English version in [Reer 94c]) for the quantification of such interactions.

5.9 Requirements for a method for human reliability assessment

In several publications, the different HRA methods are discussed and often criticised according to the deficiencies of the methods. Some of these are:

- The problem of the completeness of the analysis. Incompleteness of the methods concerning organisational effects or cognitive errors may result in an erroneous estimation of the total risk, for instance ([Reer 93b], [Bley 92]).
- The problem of insufficient data for assessment and the problem of the relevance of the collected data. HRA data is currently not available for most of the situations to be assessed. Therefore, the necessary transmissions and judgements decrease the accuracy of the used data. This is only covered by high uncertainty bounds. This problem is intensified by the fact that no systematic investigation of the validity of the methods has been performed ([Sträter 94b], [ACSNI 91], [Fujita 92]).

- The problem of modelling human failures in PSA. Firstly, the modelling of human errors into FTA (fault tree analysis) and ETA (error tree analysis) is a simplified assumption treating the human error like a component error. This problem also includes the fact that the probabilistic quantification is to some extent insufficient for depicting the complexity of human behaviour ([Reichart 92], [Heslinga 93]). Secondly, concerning the treatment of PSF (Performance Shaping Factors), the methods only provide insufficient information about the coherency of different PSFs and a given situation. Due to this problem, the analysis is incomplete as is the process for optimisation ([Hollnagel 92]). Thirdly, concerning the error types, cognitive errors are insufficiently considered, but are of great importance particularly for the assessment of human reliability in complex disturbance situations especially in computerised technologies and modified types of organisation of work ([Reichart 92], [Reer 93b]).

To structure the different problems and enable an assessment of the methods according to the criticism, a catalogue of requirements will be introduced. These requirements include aspects concerning the methods and the information used by the methods. The catalogue will be subdivided according to the following categories of human reliability assessment:

1. Objectivity of the method
2. Validity of the method
3. Reliability of the method

These distinctions are used in different scientific disciplines like mathematics, philosophy, psychology and can also be considered to subdivide requirements for human reliability assessment (see [Bubb 92], [ACSNI 92], [Swain 89]).

In the context of the assessment of operator actions, objectivity of a method means that the method is able to measure real human failures in the framework of PRA. If a method is less objective, it measures less human failures, but more technical failures for instance. In the theoretical scientific sense, objectivity means precision about what a human mistake is. It is to be distinguished from objectivity in the sense of 'opposite to subjectivity'. In the mathematical sense, objectivity means the correlation between the parameter measured by the method and the real object which should be measured.

Validity is defined as the power of evidence of the method. It means that a method is able to measure the parameters which are used by the method. In the mathematical sense, validity means the correlation of the measured variable with the parameter the method intends to measure.

Reliability of the method is the quality of a method to repeat gained predictions again at a later time or by different persons. In the mathematical sense, reliability means the correlation between the variable measured at different times or by different persons.

Suppose a method for human reliability assessment is optimal in the modelling of personal actions, but the underlying data are insufficient. This means that the method would be an objective one, but would be less valid. Objectivity, validity and reliability are not independent of each other and rely on each other. For an optimal HRA method a maximum of objectivity, validity and reliability is required.

5.9.1 Catalogue of requirements

Starting from the results of a literature review concerning actual methods for HRA, a catalogue of requirements for HRA methods is introduced according to the above-mentioned categories of requirements. The catalogue may be used for assessment and for the comparison of applications of HRA methods as well as for the evaluation of a new method. The catalogue of requirements is based on requirements mentioned in [IAEA 92], [Bubb 92], [ACSNI 91] and [Swain 89].

Because the three categories of objectivity, validity and reliability are used, this catalogue of requirements may be more balanced with respect to the different requirements than other classifications which only focus on parts of this catalogue. [Swain 89] mainly focuses on the reliability and [ACSNI 91] mainly considers the validity of the methods.

5.9.1.1 Objectivity of the method

Relevance of the Results:

The method has to provide (1) quantitative results for the operator action to be considered in PSA and (2) qualitative results for improvement of the safety and availability of the investigated plant.

Completeness of the Method:

The method has to be complete regarding (1) all types of human actions, (2) the necessary degree of detail of an action, (3) the different depths in coarse, detailed and sensitivity analysis, (4) the error types to be analysed, (5) all factors influencing the action and (6) the definition of an operator error.

5.9.1.2 Validity of the method

Data (qualitative and quantitative):

(1) Original data of the investigated sector of industry should be used (i.e., nuclear power plants in the nuclear area). (2) Data from simulators may be used, if plant experience is lacking (e.g., in the case of accident management measures). The transferability of the simulator experiments has to be proved by checking the circumstances of the experiments. (3) Other data (non-nuclear industry, field studies, laboratory experiments, expert judgement) may be used, if neither nuclear plant experience nor simulator experiments are available. The transferability of the data has to be approved. Expert judgement has to be performed by considering psychological constraints of human judgements.

Approval of the Validity of the Method:

Regarding the validity of the assessment the following is to be demanded:

- the method must be able to describe human actions realistically
- the data on human behaviour, on which the HRA method was based, has to be transferable to the application of the HRA method in a particular PSA study.
- indications of the validity of the results like verification by plant experience or simulator experiments

- convergence, i.e. the results have to be in accordance with the results of other methods which have to fit this catalogue of requirements as well
- compatibility of the method, the underlying assumptions and the constraints with human factor knowledge as manifested in widely recognised publications
- common acceptance of the method, criteria are, for instance: the use of the method in various safety studies and by various users, review of the method by other experts and accommodation of the methods according to qualified criticism on the method.

5.9.1.3 *Reliability of the method*

Consistency of the Results:

If the method is used for a given task (1) by different users or, (2) by the same user at different times the results have to agree sufficiently.

Accessibility to the Information:

The information in which the method is described and the needed data are documented has to be public available and transparently documented.

User Aids:

Within the method, appropriate, well-defined and understandable rules for application have to be available. If subjective judgement by user is required, the judgement process has to be supported in a qualified manner.

Proofing:

All assumptions, models and data, which are underlying the method, have to be provable and testable (cf. accessibility to the information). A user has to be guided in a way that his results are provable.

5.9.2 *Summary of the catalogue of requirements*

The catalogue of requirements was structured according to the criteria of objectivity, validity and reliability. It allows the weak and strong points of different HRA methods to be worked out as well as a new HRA method to be evaluated. Table 5-26 summarises the catalogue in a short overview.

Table 5-26. Summary of the catalogue of requirements.

<i>Criteria</i>	<i>Requirements of a method for the assessment of human reliability</i>
Objectivity of the method	
Relevance of the results	Does the method provide quantitative and qualitative results for PSA and plant safety?
Completeness of the method	Are all types of human actions considered?
	Is the required degree of detail of an action considered?
	Are coarse, detailed, and sensitivity-analysis executed?
	Are error types considered completely and is an error definition given?
	Are the effective influencing factors considered completely?
Validity of the method	
Data (qualitative and quantitative)	Is the following ranking of data sources considered? 1. Original data from the relevant production area (e.g., NPP) 2. NPP simulators (+ description of constraints) 3. Non-nuclear industry, field studies, laboratory experiments, (+ careful review of the transferability) 4. Expert judgement (+ acknowledged method for judgement)
Proof of validity	Are the human actions modelled accurately?
	Are the used data on human errors transferable to the actions to be inspected?
	Can the validity of the results be proved?
	Is the method in concordance with results of other qualified methods?
	Is the method in concordance with human factor knowledge?
	Is the method acknowledged by other experts?
Reliability of the method	
Consistency of the results	Does concordance exist between different users?
	Does concordance exist in the results by using the method at various times?
Accessibility of the information	Is the information generally available?
	Is the information transparent?
User aids	Are application rules or application instruction available?
	Will subjective ratings be supported in a qualified manner?
Proofing	Are the underlying assumptions, models and data provable and testable?

5.10 Conclusions

The variety of methods and the differences in their results show that the field of predictive human reliability analysis (HRA) remains a young field.

The current version of THERP (published 1983) is an established and widely accepted tool. However, THERP has some weak points, especially with regard to guidance and data for the quantification of cognitive error mechanisms in decision-making. Some new developments (published 1984-1994) offer approaches to diminish these weak points:

- data-based time-reliability correlations (TRCs) (EdF's PHRA, HCR/ORE),
- extended guidelines and data for the incorporation of PSFs (SLIM, HEART, INTENT).

On the whole, THERP is still a standard framework for HRA. The new developments (outlined above) make use of essential basic concepts and data from THERP, e.g.:

- breaking down an accident response into diagnosis and post-diagnosis actions,
- using TRCs for the quantification of diagnosis failures,
- dependence model,
- HEP adjustment according to PSFs, or
- some specific HEPs from the THERP data tables (see Figure A-6).

In [Reer 93b] 21 weak points of the current HRA approach are identified. In addition to data acquisition, there is still a need for basic investigation, especially on the subjects of **dependence** (see Section 5.6.6) and **time** (Section 5.7.4).

5.11 Chapter References

- /1/ [ACSNI 91] Advisory Committee on the Safety of Nuclear Installations, *Human Reliability Assessment – a Critical Overview*. Study Group on Human Factors. Second Report. HMSO. London (GB), 1991
- /2/ [Barriere 94] M. Barriere, W. Luckas, D. Whitehead, A. Ramey-Smith, *An Analysis of Operational Experience During Low Power and Shutdown and a Plan for Addressing Human Reliability Assessment Issues*. NUREG/CR-6093, Washington, DC (USA), 1994
- /3/ [Beare 84] A. N. Beare, R. S. Dorris, C. R. Bovell, D. S. Crowe, E. J. Kozinsky, *A Simulator-Based Study of Human Errors in Nuclear Power Plant Control Room Tasks*. NUREG/CR-3309, Washington, DC (USA), 1984
- /4/ [Berg 92] H. P. Berg, H. Schott, *Stand von Wissenschaft und Technik auf dem Gebiet der Quantifizierung der menschlichen Zuverlässigkeit – Dezember 1991* –. Bundesamt für Strahlenschutz, Fachbereich Kerntechnische Sicherheit, KT-2/92, Salzgitter (D), 1992

- /5/ [Bello 80] G. C. Bello, V. Colombari, The Human Factors in Risk Analysis of Process Plants: The Control Room Model, TESEO. *6th Advances in Reliability Technology Symposium*. NCSR-R23, United Kingdom Atomic Energy Authority, Warrington (GB), 1980
- /6/ [Bley 92] D. Bley, S. Kaplan, D. Johnson, The Strengths and Limitations of PSA: Where We Stand, *Reliability Engineering and System Safety* **38** (1992) pp. 3–26
- /7/ [BMI 77] Besondere Vorfälle in Kernkraftwerken der Bundesrepublik Deutschland. Berichtszeitraum: 1965-1976. Der Bundesminister des Innern, Bonn (D), 1977
- /8/ [Bubb 92] H. Bubb (Ed.), Menschliche Zuverlässigkeit. EcoMed, Landsberg (D), 1992.
- /9/ [Bubb 93] H. Bubb, H. Schmidtke, Systemergonomie. In: H. Schmidtke (Ed.), Ergonomie. Hanser, München (D), 1993
- /10/ Carnino 95] A. Carnino, G. Weimann (Eds.), Proceedings of the International Topical Meeting on Safety Culture in Nuclear Installations, Vienna, 24.-25. April 1995. American Nuclear Society, Austria Local Section, Vienna (A), 1995
- /11/ Chien 88] Chien, S.H., A.A. Dykes, J.W. Stetkar, D.C. Bley, “Quantification of Human Error Rates Using a SLIM-Based Approach,” 1988 IEEE Fourth Conference on Human Factors and Power Plants, Monterey, CA, June 5-9, 1988.
- /12/ [Coombs 75] C. H. Coombs, R. M. Dawes, A. Tversky, Mathematische Psychologie. Beltz, Weinheim (D), Basel (CH), 1975
- /13/ [Dang 95] Dang V.N., and Hirschberg S. Human Reliability Analysis in Probabilistic Safety Assessments: Current issues, the Swiss studies and options for research. HSK-AN-2887, Paul Scherrer Institute for the Swiss Federal Nuclear Inspectorate, December 1995.
- /14/ [Degen 94] G. Degen, J. Mertens, B. Reer, ESAP, an Easy-To-Use Expert System for the Systematic Analysis of Operator Actions within the Scope of Probabilistic Safety Assessment. Proceedings of Seminar/Workshop on Methodological Issues of PSA/HRA and Supporting Software Tools, Gdansk, June 13-17, 1994, Gdansk (PL), 1994
- /15/ [Dougherty 93] E. M. Dougherty jr., Context and Human Reliability Analysis. *Reliability Engineering & System Safety* **41** (1993), pp. 25-47
- /16/ [Dougherty 94] E. M. Dougherty jr., A Note on HRA at PSAM-II. *Reliability Engineering & System Safety* **46** (1994), pp. 291-294
- /17/ [Drouin 87] M. T. Drouin, J. L. La Chance, B. J. Shapiro, F. T. Harper, T. A. Wheeler, Analysis of Core Damage Frequency from Internal Events: Grand Gulf, Unit 1. NUREG/CR-4550 Vol. 6, Washington, DC (USA), 1987
- /18/ [DRS-B] Deutsche Risikostudie Kernkraftwerke, Phase B. Gesellschaft für Reaktorsicherheit, Köln und Garching. Verlag TÜV Rheinland, Köln (D), 1990.

- /19/ [Edwards 77] W. Edwards, How to Use Multi-Attribute Utility Measurement for Social Decision Making. IEEE Transactions on Systems, Man and Cybernetics, SMC-7, 5, 1977.
- /20/ [Embrey 83] D. E. Embrey, The Use of Performance Shaping Factors and Quantified Expert Judgement in the Evaluation of Human Reliability: An Initial Appraisal. NUREG/CR-2986, Washington, DC (USA), 1983
- /21/ [Embrey 84] D. E. Embrey, P. Humphreys, E. A. Rosa, B. Kirwan, K. Rea, SLIM-MAUD: An Approach to Assessing Human Error Probabilities Using Structured Expert Judgement, Vol. I: Overview of SLIM-MAUD, Vol. II: Detailed Analyses of the Technical Issues. NUREG/CR-3518, Washington, DC (USA), 1984
- /22/ [Embrey 86] D. E. Embrey, J. T. Reason, The Application of Cognitive Models to the Evaluation and Prediction of Human Reliability. Proceedings of the International Meeting on Advances in Human Factors in Nuclear Power Systems, Knoxville, 21-24 April 1986. American Nuclear Society, LaGrange Park (USA), 1986
- /23/ [ESCIS 86] Einführung in die Risikoanalyse. Systematik und Methoden. Chemische Rundschau, Separatdruck, Schriftenreihe der Expertenkommission für Sicherheit in der chemischen Industrie der Schweiz (ESCIS), Heft 4, 2. unveränderte Auflage, 1986
- /24/ [EPS 900] IPSN, Framatome, EPS 900. A Probabilistic Safety Assessment of the Standard French 900 MWe Pressurized Water Reactor. Main Report. Raven International France, Clamart (F), 1990
- /25/ [EPS 1300] EPS 1300. Probabilistic Safety Assessment of Reactor Unit 3 in the Paluel Nuclear Power Centre (1300 MWE). Overall Report. Electricité de France, 1990
- /26/ [ESI 92] ESI VIII, Risk Management in Complex Production and Transportation System. Final Report, Eighth European Summer Institute (ESI VIII), Ed.: Swedish OR Association, c/o FOA, Sundbyberg (S), 1992
- /27/ [Flanagan 54] J. C. Flanagan, The Critical Incident Technique. Psychol. Bull. 51 (1954), pp. 327-358
- /28/ [Fleming 75] K. N. Fleming, P. H. Raabe, G. W. Hannaman, W. J. Houghton, R. D. Pfremmer, F. S. Dombek, HTGR Accident Initiation and Progression Analysis Status Report, Volume II, AIPA Risk Assessment Methodology. GA/A13617 Vol. II UC-77, General Atomic Co., San Diego (USA), 1975
- /29/ [Fujita 92] Y. Fujita, Human Reliability Analysis - A Human Point of View. Reliability Engineering and System Safety 38 (1992), pp. 71-79
- /30/ [Gautschi 89] K. Gautschi, Ausbildung und Übungen des Notfallstabs der Picketingenieure und der Notfallequipen. In: [SVA 89]
- /31/ [Gerdes 93] V. G. J. Gerdes, HRA Techniques; A Selection Matrix. KEMA, Risk and Reliability Analysis group, the Netherlands. Proceedings of the SRE symposium, Arnhem, The Netherlands, October 1993: 12 pp

- /32/ [Gertman 92] D. I. Gertman, H. S. Blackmann, L. N. Haney, K. S. Seidler, H. A. Hahn, INTENT: A Method for Estimating Human Error Probabilities for Decision Based Errors. Reliability Engineering & System Safety 35 (1992) pp.127-136
- /33/ [Ghertman 85] F. Ghertman, P. Dietz, Human Error Data Collection Analysis Program Undertaken since 1982 by Electricité de France with INPO. Proceedings of the ANS/ENS Topical Meeting on Probabilistic Safety Methods and Applications, Vol. 2, Paper 89. San Francisco (USA), 1985
- /34/ [Grote 93] G. Grote, C. Künzler, Sicherheit in soziotechnischen Systemen. Zwischenbericht. ETH Zürich, Polyprojekt „Risiko und Sicherheit technischer Systeme“, Polyprojekt-Bericht 05/93, Zürich (CH), 1993
- /35/ [Haas 82] P. M. Haas, T. F. Bott, Criterion for Safety Related Plant Operator Actions: A Preliminary Assessment of Available Data. Reliability Engineering 3 (1982) pp.59-72
- /36/ [Hacker 80] W. Hacker, Allgemeine Arbeits- und Ingenieurpsychologie. VEB Deutscher Verlag der Wissenschaften, Berlin (D), 1980
- /37/ [Hacker 86] W. Hacker, Arbeitspsychologie. Huber, Bern (CH), 1986
- /38/ [Hall 82] R. E. Hall, J. Fragola, J. Wreathall, Post Event Human Decision Errors: Operator Action Tree / Time Reliability Correlation. NUREG/CR-3010, Washington, DC (USA), 1982
- /39/ [Hannaman 84] G. W. Hannaman, A. J. Spurgin, Y. Lukic, Human Cognitive Reliability Model for PRA Analysis. Draft report NUS-4531, EPRI Project RP2170-3, San Diego (USA), 1984
- /40/ [Hannaman 85] G. W. Hannaman, A. J. Spurgin, Y. Lukic, A Model for Assessing Human Cognitive Reliability in PRA studies. IEEE Third Conference on Human Factors in Nuclear Power Plants, Monterey, California, June 23-27, 1985. Institute of Electronic and Electrical Engineers, New York (USA), 1985
- /41/ [Hannaman 86] G. W. Hannaman, P. Moieni, J. P. Spurgin, La Salle Human Reliability Measurement Program Data Analysis Report. Draft NUS-4965, NUS Corporation, San Diego, Nov. 1986, (Proprietary)
- /42/ [Hannaman 88] G. W. Hannaman, D. H. Worledge, Some Development in Human Reliability Analysis Approaches and Tools. Reliability Engineering & System Safety 22 (1988) pp.235-257
- /43/ [Hansmann 89] W. Hansmann, Notfallübungen, Empfehlungen und Mitarbeit der Aufsichtsbehörden. In: [SVA 89]
- /44/ [Hennings 95] W. Hennings, J. Mertens, B. Reer, Methodik der Risikoanalyse für Kernkraftwerke. Eine bewertende Bestandsaufnahme mit Blick auf regionale Sicherheitsplanung. vdf Verlag der Fachvereine, Zürich (CH), 1995

- /45/ [von Herrmann 83a] J. L. von Herrmann, Methods for Review and Evaluation of Emergency Procedure Guidelines, Volume I: Methodologies. NUREG/CR-3177, Washington, DC (USA), 1983
- /46/ [von Herrmann 83b] J. L. von Herrmann, W. A. Brinsfield, R. E. Brown, Methods for Review and Evaluation of Emergency Procedure Guidelines, Volume II: Applications to Westinghouse Plants. NUREG/CR-3177, Washington, DC (USA), 1983
- /47/ [Heslinga 93] G. Heslinga, H. Arnold, Human Reliability: To what extent can we consider Humans as System Components. Proceedings, ENS TOPNUX 1993. Hague, 25-28 April 1993, The Netherlands
- /48/ [Hirschberg 90] S. Hirschberg (Ed.), Dependencies, Human Interactions and Uncertainties in Probabilistic Safety Assessment. Final Report of the NKA Project RAS 470, Västerås, Sweden. ABB Atom AB Library, prepared by ABB Atom, Sweden, Risø National Laboratory, Denmark, Studsvik AB, Sweden, Technical Research Centre of Finland, 1990
- /49/ [Hoffmann 89] E. Hoffmann, Vorbereitung am Simulator zur Handhabung auch schwerer Störfälle. In: [SVA 89]
- /50/ [Hoffmeister 74] N. Hoffmeister, Analyse einiger wichtiger Vorkommnisse in Kernkraftwerken. In: SVA (Ed.), Informationstagung über die Sicherheit von Kernkraftwerken, 25./26. November 1974. Tagungsreferate. SVA, Zürich (CH), Bern (CH), 1975
- /51/ [Hollnagel 92] E. Hollnagel, The Reliability of Man-Machine Interaction. Reliability Engineering and System Safety 38 (1992) p.81 ff.
- /52/ [IAEA 89] Models and Data Requirements for Human Reliability Analysis. IAEA TECDOC 499. IAEA, Vienna (A), 1989
- /53/ [IAEA 90] Human Error Classification and Data Collection. IAEA TECDOC 538. IAEA, Vienna (A), 1990
- /54/ ([IAEA 92] Procedures for Conducting Probabilistic Safety Assessments of Nuclear Power Plants (Level 1). Safety Series No. 50 P, IAEA, Vienna (A), 1992, p.51ff
- /55/ [Indian Point 82] Indian Point Probabilistic Study. Power Authority of the State New York and Consolidated Edison Company of New York, Inc., New York (USA), 1982
- /56/ [IRS 72] IRS Kurzinformationen. Kernkraftwerk Obrigheim, Bersten eines Entwässerungsbehälters durch Kühlmittelverlust über das Entwässerungssystem. Institut für Reaktorsicherheit der Technischen Überwachungsvereine (IRS), Köln (D), 1972
- /57/ [Kemeny 79] S. G. Kemeny, Report of the President's Commission on the Accident at Three Mile Island. U.S. Government Printing Office, Washington, DC (USA), 1979
- /58/ [KFA 81] KFA, GRS, Sicherheitsstudie für HTR-Konzepte unter deutschen Standortbedingungen, Hauptband zur Phase I B. Spezielle Berichte der Kernforschungsanlage Jülich - Nr. 136, KFA, Jül-Spez-136 Bd. 1, Jülich (D), 1981

- /59/ [KFA 84] KFA, Sicherheitstechnische Untersuchungen zum Störfallverhalten des HTR-500. Spezielle Berichte der Kernforschungsanlage Jülich - Nr. 240, KFA, Jül-Spez-240, Jülich (D), 1984
- /60/ [Kirwan 94] B. Kirwan, private note, 1994
- /61/ [Kirwan 95] B. Kirwan, private note, 1995
- /62/ [Kolaczowski 86] A. M. Kolaczowski, J. A. Lambright, W. L. Ferrell, N. G. Cathey, B. Najafi, F. T. Harper, Analysis of Core Damage Frequency from Internal Events: Peach Bottom, Unit 2. NUREG/CR-4550 Vol. 4, Washington, DC (USA), 1986
- /63/ [Kolb 82] G. J. Kolb, D. M. Kunsman, B. J. Bell, N. L. Brisbin, D. D. Carlson, S. W. Hatch, D. P. Miller, B. J. Roscoe, D. W. Stack, R. B. Worell, J. Robertson, R. O. Wooton, S. H. McAhren, W. L. Ferrell, W. J. Galyean, J. A. Murphy, Interim Reliability Program: Analysis of the Arkansas Nuclear One - Unit 1 Nuclear Power Plant. NUREG/CR-2787 Vols. 1 and 2, Washington, DC (USA), 1982
- /64/ [Kosmowski 94] K. T. Kosmowski, G. Degen, J. Mertens, B. Reer, Development of Advanced Methods and Related Software for Human Reliability Evaluation within Probabilistic Safety Analyses. Berichte des Forschungszentrums Jülich; 2928, KFA, Jül-2928, Jülich (D), 1994
- /65/ [Kosmowski 94b] K. T. Kosmowski, private note, 1994
- /66/ [Kozinski 84] E. J. Kozinski, L. H. Grey, A. N. Beare, D. B. Burks, F. E. Gomer, Safety-Related Operator Actions: Methodology for Developing Criteria. NUREG/CR-3515, Washington, DC (USA), 1984
- /67/ [Lanore 87] J. M. Lanore, J. L. Caron, A. Ellia-Hervy, J. L'Henoret, Interaction Between Thermal/Hydraulics, Human Factors and System Analysis for Assessing Feed and Bleed Risk Benefits. In: Probabilistic Safety Assessment and Risk Management/PSA'87, Volume 1, pp 116-119. Europ. Nuclear Soc. / Swiss Nuclear Soc., Verlag TÜV Rheinland, Köln (D), 1987
- /68/ [Madjar 93] M. Madjar, Überblick über Methoden und Einflußgrößen bei der Risikoermittlung chemischer Anlagen. ETH Zürich, Polyprojekt „Risiko und Sicherheit technischer Systeme“, Polyprojekt-Bericht 04/93, Zürich (CH), 1993
- /69/ [Miller 87] D. P. Miller, A. D. Swain, Human Error and Human Reliability. In: G. Salvendy (Ed.), Handbook of Human Factors/Ergonomics, Chapter 2.8. Wiley & Sons, New York (USA), 1987
- /70/ [Moieni 86] P. Moieni, G. W. Hannaman, Uncertainty Analysis in Time-Dependent Human Reliability Models - Application to the Human Cognitive Reliability (HCR) Model. NUS-4915, NUS Corporation, Gaithersburg, MD (USA), 1986
- /71/ [Moieni 94] P. Moieni, J. Spurgin, A. Singh, Advances in Human Reliability Analysis Methodology. Part I: Frameworks, Models and Data. Reliability Engineering and System Safety 44 (1994), pp. 27-55

- /72/ [Mosneron 89] F. Mosneron Dupin, Méthode pratique de prise en compte du Facteur Humain dans les études de séquences accidentelles. EPS FH 006 c, Electricité de France, 1989, (restricted circulation)
- /73/ [Mosneron 90] F. Mosneron Dupin, A. Villemeur, J. M. Moroni, Paluel Nuclear Power Plant PSA: Methodology for Assessing Human Reliability. 7th International Conference on Reliability and Maintainability, Brest, 1990, Brest (F), 1990
- /74/ [Mosneron 92] F. Mosneron Dupin, G. Saliou, F. Lars, Probabilistic Human Reliability Analysis: The Lessons Derived from Plant Operation at Electricité de France. In: IAEA (Ed.), Use of Probabilistic Safety Assessment for Operational Safety. PSA '91. Proceedings of an International Symposium, Vienna, June 3-7, 1991. IAEA-SM-321/57, Vienna (A), 1992
- /75/ [Mosneron 93] F. Mosneron Dupin, Characteristics of the EdF Probabilistic Human Reliability Analysis Methodology. Note Interne Technique, T54/93-20, EDF/DER/RNE, Electricité de France, 1993
- /76/ [Mosneron 94] F. Mosneron Dupin, Is Probabilistic Human Reliability Analysis Possible? Presented at the EdF International Seminar on PSA and HRA, Paris, November 21-23, 1994, Paris (F), 1994
- /77/ [Mosneron 95] F. Mosneron Dupin, private note, 1995
- /78/ [Mosneron 96] F. Mosneron Dupin, private note, 1996
- /79/ [MUSA] D. Haschke, Die probabilistische Sicherheitsanalyse des Kernkraftwerks Mühleberg. In: SVA-Vertiefungskurs „Fortgeschrittene Sicherheitsanalyse“. Schweizerische Vereinigung für Atomenergie, Bern (CH), 1991.
- /80/ [Musicki 94] Musicki, Z., Chu, T.L., Ho V., Hou, Y.-M., Lin, J., Yang, J., Siu, N., Evaluation of Potential Severe Accidents During Low Power and Shutdown Operations at Surry, Unit 1: Analysis of Core Damage Frequency from Internal Fires During Mid-Loop Operations, Main Report, NUREG/CR-6144, U.S. Nuclear Regulatory Commission, July 1994.
- /81/ [NUREG-1154] Loss of Main and Auxiliary Feed Water Event at the Davis-Besse Plant on June 9, 1985. NUREG-1154, Washington, DC (USA), 1985
- /82/ [Oswald 82] A. J. Oswald, C. D. Gentillon, S. D. Matthers, T. R. Mitchum, Generic Data Base for Data and Models Chapter of the National Reliability Evaluation Program (NREP) Guide. EGG-EA-5887 Informal Report, EG&G Idaho, Inc., Idaho Falls (USA), 1982
- /83/ [Parry 91] G. W. Parry, A. Singh, A. Spurgin, P. Moieni, A. Beare, An Approach to the Analysis of Operating Crew Responses With Simulator Exercises. In: U. Hauptmanns (Ed.), Proceedings of the OECD/BMU Workshop on Special Issues of Level 1 PSA. GRS-86, Köln (D), 1991
- /84/ [Parry 92] G. W. Parry et al., An Approach on the Analysis of Operator Actions in Probabilistic Risk Assessment. EPRI TR-100259, Electric Power Research Institute, Palo Alto (USA), 1992

- /85/ [Pew 81] R. W. Pew, D. C. Miller, C. E. Feeher, Evaluation of Proposed Control Room Improvements Through Analysis of Critical Operator Decisions. EPRI-NP-1082, Electric Power Research Institute, Palo Alto (USA), 1981
- /86/ [Phillips 85] L. D. Phillips, P. Humphreys, D. E. Embrey, D. L. Selby, A Socio-Technical Approach to Assessing Human Reliability (STHR). In: Pressurized Thermal Shock Evaluation of the Calvert Cliffs Unit 1 Nuclear Power Plant, Appendix D. NUREG/CR-4022, Washington, DC (USA), 1985
- /87/ [Potash 81] L. M. Potash, M. Stewart, P. E. Dietz, C. M. Lewis, E. M. Dougherty Jr., Experience in Integrating the Operator Contributions in the PRA in Actual Operating Plants. In: Proceedings of the ANS/ENS Topical Meeting on Probabilistic Risk Assessment, Port Chester, NY, September 1981, Vol. II, pp. 1054-1063. American Nuclear Society, LaGrange Park (USA), 1981
- /88/ [Poucet 88] A. Poucet, Survey of Methods Used to Assess Human Reliability in the Human Factors Reliability Benchmark Exercise. Reliability Engineering & System Safety 22 (1988) pp. 257-268
- /89/ [Poucet 89] Poucet, A., "The European Benchmark Exercise on Human Reliability Analysis," Proc. of the Int'l Topical Meeting on Probability, Reliability, and Safety Assessment, PSA '89, Pittsburgh, PA, U.S.A., April 2-7, 1989.
- /90/ [PRA-PG] PRA Procedures Guide – A Guide to the Performance of Probabilistic Risk Assessments for Nuclear Power Plants. Final Report. NUREG/CR-2300, Washington, DC (USA), Jan. 1983
- /91/ [Rasmussen 79] J. Rasmussen, On the Structure of Knowledge - A Morphology of Mental Models in a Man-Machine Context. Risø-M-2192, Risø National Laboratory, Roskilde (DK), 1979
- /92/ [Reason 90] J. Reason, Human Error. Cambridge University Press, Cambridge (GB), 1990
- /93/ [Reason 94] J. Reason, Errors, Outcomes and Circumventions: a Reply to Dougherty, Reliability Engineering & System Safety 46 (1994), pp. 297-298
- /94/ [Reer 88] B. Reer, Deutsche Übersetzungen und Erläuterungen zu den Daten aus dem Swain-Handbuch über menschliche Zuverlässigkeit in Kernkraftwerken. Interner Bericht, KFA-ISF-IB-6/88, Forschungszentrum Jülich GmbH, Jülich (D), 1988
- /95/ [Reer 93] B. Reer, Entscheidungsfehler des Betriebspersonals von Kernkraftwerken als Objekt probabilistischer Risikoanalysen. Berichte des Forschungszentrums Jülich; 2811, KFA, Jül-2811, Jülich (D), 1993
- /96/ [Reer 93b] B. Reer, J. Mertens, Zukünftige Forschungsthemen zur systematischen Erweiterung der Methodik probabilistischer Mensch-Maschine Systemanalysen. Interner Bericht, KFA-ISR-IB-9/93. Institut für Sicherheitsforschung und Reaktortechnik (ISR), Forschungszentrum Jülich GmbH (KFA), Jülich (D), 1993

- /98/ [Reer 94a] B. Reer, Dynamische Analysemethode für Operateurhandlungen (DAO): Ein Verfahren zur Quantifizierung der zeitabhängigen Zuverlässigkeit von Notfallmaßnahmen. In: Deutsches Atomforum e.V. und Kerntechnische Gesellschaft e.V. (Ed.), Jahrestagung Kerntechnik '94. Tagungsbericht. INFORUM Verlagsgesellschaft, Bonn (D), 1994
- /99/ [Reer 94b] B. Reer, A Probabilistic Method for Analysing the Reliability Effect of Time and Organisational Factors. European Journal of Operational Research 75 (1994), pp. 521-539. (Revised version of a paper presented in [ESI 92])
- /100/ [Reer 94c] B. Reer, Incorporation of Risk-Taking Behaviour into Human Reliability Analysis. In: A. Stritar (Ed.), International Meeting: PSA/PRA and Severe Accidents '94. Nuclear Society of Slovenia, Ljubljana (SLO), 1994
- /101/ [Reer 94d] B. Reer, F. Przybylski, W. Ullwer, An Extended View of Plant Specific Data for Human Reliability Analysis: The Impact of Failure Data and Modelling Details. In: H.-P. Balfanz (Ed.), 4th TÜV Workshop on Living PSA Application in Hamburg, 2-3 May 1994. TÜV Nord e.V., Hamburg (D), 1994
- /102/ [Reer 96] Reer, B., Sträter, O. & Mertens, J. (1996) Evaluation of Human Reliability Methods Addressing Cognitive Error Modelling and Quantification. Jülich; 3222. ISSN 0944-2952. KFA-Jülich. Jülich.
- /103/ [Reichart 85] G. Reichart, in: W. Frey, H. Hörtner, J. von Linden, G. Rappl, G. Reichart, Deutsche Precursor Studie - Auswertung anlagenspezifischer Betriebserfahrung im Hinblick auf Vorläufer zu möglichen schweren Kernschäden. Report GRS-A-1149, GRS, Köln (D), Garching (D), 1985
- /104/ [Reichart 92] G. Reichart, in: H. Bubb (Ed.), Menschliche Zuverlässigkeit. EcoMed, Landsberg (D), 1992, pp. 106 ff.
- /105/ [Roth-Seefried 89] H. Roth-Seefried, Betriebsanweisungen zur Störfallbeherrschung bei Siemens/KWU-Druckwasser-Reaktoren. In: [SVA 89]
- /106/ [Samanta 85] P. K. Samanta, J. N. O'Brien, H. W. Morrison, Multiple Sequential Failure Model: Evaluation of and Procedure for Human Error Dependency. NUREG/CR-3837, Washington, DC (USA), 1985
- /107/ [Seaver 83] D. A. Seaver, W. G. Stillwell, Procedures for Using Expert Judgement to Estimate Human Error Probabilities in Nuclear Power Plant Operations. NUREG/CR-2743, Washington, DC (USA), 1983
- /108/ [Semmer 94] N. Semmer, Der menschliche Faktor in der Arbeitssicherheit: Mechanismen, Verhütung und Korrektur von menschlichen Fehlhandlungen. In: SVA (Ed.), SVA-Ausbildungsseminar (Workshop), Ursachenanalyse von Störfällen in Kernkraftwerken. Schweizerische Vereinigung für Atomenergie (SVA), Bern (CH), 1994
- /109/ [SHARP] G. W. Hannaman, A. J. Spurgin, Systematic Human Action Reliability Procedure (SHARP). EPRI-NP-3583, Electric Power Research Institute, Palo Alto (USA), 1984

- /110/ [Siegel 84] A. I. Siegel, W. D. Bartter, J. J. Wolf, H. E. Knee, Maintenance Personnel Performance Simulation (MAPPS) Model: Description of Model Content, Structure, and Sensitivity Testing. NUREG/CR-3626, Washington, DC (USA), 1984
- /111/ [Smidt 79] D. Smidt, Reaktorsicherheitstechnik. Springer, Berlin (D), 1979
- /112/ [Spurgin 90] A. J. Spurgin et al., Operator Reliability Experiments Using Power Plant Simulators, Vols. 1 and 2: Executive Summary and Technical Report. EPRI-NP-6937, Electric Power Research Institute, Palo Alto (USA), 1990
- /113/ [Sträter 94a] O. Sträter, The Role of Plant Experience to Consider the Human Factor in Living PSA. In: H.-P. Balfanz (Ed.), 4th TÜV-Workshop on Living PSA Application in Hamburg, 2.-3. May 1994. TÜV Nord, Hamburg (D), 1994
- /114/ [Sträter 94b] O. Sträter, W. Preischl, A. Berning, Vergleichende Auswertung probabilistischer Sicherheitsanalysen im Bereich Personalhandlungen: Untersuchung spezieller Methodenfragen zu Personalhandlungen. GRS-A-2151, Gesellschaft für Reaktorsicherheit (GRS) mbH, Köln (D), 1994
- /115/ [Sträter 95a] O. Sträter, Eine Methode zur Erfassung und Auswertung von Betriebserfahrung im Hinblick auf menschliche Fehler. In: Technische Zuverlässigkeit. Tagung Fulda. 26/27.9.1995, VDI-GSP, Düsseldorf (D), 1995
- /116/ [Sträter 95b] O. Sträter, Ergonomic Principles for Accident Management Support (AMS). Final Report for the EU project on AMS, Gesellschaft für Reaktorsicherheit (GRS) mbH, Köln (D), 1995
- /117/ [SVA 89] SVA-Vertiefungskurs „Störfallmanagement in Kernkraftwerken“. Schweizerische Vereinigung für Atomenergie, Bern (CH), 1989
- /118/ [Swain 80] A. D. Swain, H. E. Guttmann, Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications. Draft Report. NUREG/CR-1278, Washington, DC (USA), 1980
- /119/ [Swain 83] A. D. Swain, H. E. Guttmann, Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications. Final Report. NUREG/CR-1278, Washington, DC (USA), 1983
- /120/ [Swain 87] A. D. Swain, Accident Sequence Evaluation Program Human Reliability Analysis Procedure. NUREG/CR-4772, Washington, DC (USA), 1987
- /121/ [Swain 88] A. D. Swain, L. M. Weston, An Approach to the Diagnosis and Misdiagnosis of Abnormal Conditions in Post-Accident Sequences in Complex Man-Machine Systems. In: L. P. Goodstein, H. Anderson, S. Olson (Ed.), Task, Errors and Mental Models. Taylor & Francis, London (GB), New York (USA), 1988
- /122/ [Swain 89] A. D. Swain, Comparative Evaluation of Methods for Human Reliability Analysis. GRS-71, Gesellschaft für Reaktorsicherheit (GRS) mbH, Köln (D), Garching (D), 1989

- /123/ [Villemeur 86] A. Villemeur, J. M. Moroni, F. Mosneron Dupin, T. Meslin, A Simulator-Based Evaluation of Operator Behaviour by Electricité de France. Proceedings of the International Meeting on Advances in Human Factors in Nuclear Power Systems, Knoxville, 21-24 April 1986. American Nuclear Society, LaGrange Park (USA), 1986
- /124/ [Vuorio 87] U. M. Vuorio, J. K. Vaurio, Advanced Human Reliability Analysis Methodology and Applications. In: Probabilistic Safety Assessment and Risk Management / PSA'87, Volume 1, pp. 104-109. Europ. Nuclear Soc. / Swiss Nuclear Soc., Verlag TÜV Rheinland, Köln (D), 1987
- /125/ [Wakefield 87] D. J. Wakefield, C. D. Adams, Quantification of Dynamic Human Errors in the TMI-1 PRA. In: Probabilistic Safety Assessment and Risk Management / PSA'87, Volume 1, pp. 110-115. Europ. Nuclear Soc. / Swiss Nuclear Soc., Verlag TÜV Rheinland, Köln (D), 1987
- /126/ [Wakefield 92] D. I. Wakefield, G. W. Parry, G. W. Hannaman, A. J. Spurgin, SHARP1 - Revised Systematic Human Action Reliability Procedure, Final Report, TR-101711, Research Project 3206-01, Electric Power Research Institute, Palo Alto (USA), 1992
- /127/ [Wechsler 52] D. Wechsler, Range of Human Capacities. Williams & Wilkins, Baltimore (USA), 1952
- /128/ [Whalley 89] S. P. Whalley, B. Kirwan, An Evaluation of Five Human Error Identification Techniques. 5th International Loss Prevention Symposium, Oslo, 1989, Oslo (N), 1989
- /129/ [Wickens 84] C. D. Wickens, Engineering Psychology and Human Performance. C. E. Merrill Publishing Company, A Bell & Howell Company. Columbus, Toronto (CDN), 1984.
- /130/ [Williams 88] J. C. Williams, A Data-Based Method for Assessing and Reducing Human Error to Improve Operational Performance. Proceedings of the IEEE 4th Conference on Human Factors in Power Plants, Monterey, California, June 6-9, 1988. Institute of Electronic and Electrical Engineers, New York (USA), 1988
- /131/ [Woods 82] D. D. Woods, J. A. Wise, L. F. Hanes, Evaluation of Safety Parameter Display Concepts, Vol. 1. EPRI-NP-2239, Electric Power Research Institute, Palo Alto (USA), 1982
- /132/ [Woods 84] D. D. Woods, Some Results on Operator Performance in Emergency Events. Institute of Chemical Engineers Symposium Series, No. 90: 21, 1984, (quoted from: [Embrey 86])
- /133/ [Woods 88] D. D. Woods, E. Roth, H. Pople jr., Modelling Human Intention Formation for Human Reliability Assessment. Reliability Engineering & System Safety 22 (1988), pp. 169-200
- /134/ [Woods 90] D. D. Woods, E. Roth, H. Pople jr. The Cognitive Environment Simulation as a Tool for Modeling Human Performance, Main Report. NUREG/CR-5213, Vol. 2, Washington, 1990

- /135/ [Wortman 78] D. B. Wortman, S. D. Duket, D. J. Seifert, R. L. Hann, G. P. Chubb, The SAINT User's Manual. AMRL-TR-77-62, Wright-Patterson Air Force Base, OH (USA), 1978
- /136/ [Yeh 86] Y.-C. Yeh, M. G. K. Evans, Model for Human Reliability and its Incorporation in a PRA. In: Proceedings of the International Topical Meeting on Advances in Human Factors in Nuclear Power Systems, Knoxville, 21-24 April 1986. American Nuclear Society, LaGrange Park (USA), 1986
- /137/ [Zimolong 90] B. Zimolong, Fehler und Zuverlässigkeit. In: C. F. Graumann et al. (Eds.), Enzyklopädie der Psychologie, Themenbereich D, Serie III, Bd. 2. Verlag für Psychologie Hogrefe, Göttingen (D), 1990
- /138/ [Zimolong 91] B. Zimolong, Empirical Evaluation of THERP, SLIM and Ranking to Estimate HEPs. Reliability Engineering and System Safety 35 (1991), pp.1-11
- /139/ [Zion 81] Zion Probabilistic Safety Study. Commonwealth Edison Co., Chicago (USA), 1981

6. RESULTS OF HRA SURVEY

This chapter discusses the survey questionnaire on the important operator actions and their treatment. Section 6.1 describes the survey, which is divided into a main part and an extension. The focus of this extension is the “detailed treatment” for some actions found to be important in many of the PSAs surveyed.

An overview of the 21 survey responses, which were submitted by 13 countries, is presented in Section 6.2. Section 6.3 summarises the important actions found in common in the surveys. Next, Section 6.4 presents an overview of the detailed treatments for BWRs and PWRs. It should be emphasised that the questionnaire responses are significantly more detailed than the summaries presented in this chapter. In many cases, information may be found on many other important actions that were not addressed in the discussion in this chapter. In addition, the responses for plant types other than BWRs and PWRs are not covered in this chapter. The survey questionnaires (main part and detailed treatment extension) as well as the responses are presented in full in the appendices.

Preliminary results of the survey were presented at PSAM-III and PSA'96 [Hirschberg et al., 1996a, 1996b].

In Section 6.5, the comments from member countries on the methods used in their studies are presented. These views are diverse and discuss the experience gained from the application of the different methods.

Section 6.6 summarises the improvements made to plant design and procedures on the basis of findings related to human performance obtained in the course of performing the PSAs.

6.1 The Survey

The survey of HRAs was conducted in two parts. Part I is the main survey covering the HRA and its results while part II is an extension of the survey, focusing on the details of the treatment of two operator actions for each light water reactor type.

The following information was covered in Part I:

- Characterisation of plant, PSA, HRA
- Other factors important to the HRA
- Methods used
- List of important actions, description, results, and importance measures
- PSA-based improvements of plant or operational practice (including procedures) related to human interactions and the HRA

It should be noted that throughout the survey emphasis is placed on Cat. C actions, i.e. on the actions required of the operators during the scenario evolution in response to the initiating events. These are also referred to as dynamic operator actions.

The results of Part I of the survey were used to identify the important contributors (among the operator actions) that were common in the PSAs surveyed. Two actions for BWRs and two actions for PWRs were selected and covered in Part II. These detailed treatments include:

- “step-by-step” documentation of the quantification, and
- additional information on sequences, initiating events, and the contribution of these initiating events

For BWRs, the operator actions selected for detailed treatment are manual depressurization and standby liquid control system actuation; for PWRs, “feed and bleed”, and either alignment for recirculation or (response to) loss of Residual Heat Removal (RHR).

1. **Main survey.** Responses were completed and evaluated for four BWRs, nine PWRs and five advanced reactors. The overview of some characteristics of the PSAs and the HRAs carried out is provided in Section 6.2.
2. **Detailed treatment (Survey extension).** Commonalities and differences in the approaches used to model these actions were examined on the basis of “traceable” descriptions that could be followed step by step, for example, from Performance Shaping Factor (PSF) ratings to the Human Error Probability (HEP) value. Whenever needed, additional details were requested to establish a comparable set of facts about each study.

The survey responses as well as the additional detailed treatments generally do not reflect the changes in the plant and in operational practices made as a consequence of the PSA. In other words, the PSA results and analyses described in the survey responses often will not reflect the current state of the plants. In some cases, the PSAs may also have been updated, e.g. with more detailed analyses.

The survey questionnaire and the supplementary questionnaire concerning the detailed treatments are shown in Appendices B and C, respectively.

6.2 Overview of Survey Responses

In total, surveys of 21 PSAs were completed by 13 countries. Table 6-1 shows the survey responses by country and lists the PSA studies for which responses were submitted. The codes for the PSA studies are explained in Table 6-2.

It should be noted that a few of the PSA studies surveyed concern plant designs rather than actual power plants. These include the SBWR, AP600, and PIUS designs. The full survey responses, which are provided in Appendix F, provide the details of the reference plants for all of the PSAs surveyed.

Table 6-1. Survey responses listing country and PSA study.

BWR	PWR	Gas-Cooled and Other
	Belgium (Doel)	
		Canada (Pick)
Finland (TVO)	Finland (Loviisa)	
	France (P900)	
	France (P1300)	
	Germany (DRS)	Germany (HTR)
Italy (SBWR)	Italy (AP600)	Italy (PIUS)
Japan (B1100)	Japan (P1100)	Japan (LMFBR)
		Korea (R.O.K.) (WS2/3/4)
Netherlands (Dod)	Netherlands (Bor)	
	Spain (Alm)	
Switzerland (Müh)	Switzerland (Bez)	
	U.K. (Siz)	
		U.S.*

* For the U.S., refer to NUREG/CR-1560, which surveys the results from the Individual Plant Examinations (IPEs). Some of this information and references are included in this report.

Table 6-2. PSA studies (abbreviations).

Code	Country	Study
Alm	Spain	Almaraz PWR
AP600	Italy	General Electric AP600 (PWR)
B1100	Japan	Standardized 1100 MWe class BWR
Bez	Switzerland	Beznau PWR
Bor	Netherlands	Borssele PWR
Dod	Netherlands	Dodewaard BWR
Doel	Belgium	Doel Units 1 and 2 PWRs
DRS	Germany	Deutsche Risikostudie-A and -B (German Risk Study)
HTR	Germany	HTR-500 (medium-sized gas-cooled pebble bed high temperature reactor)
LMFBR	Japan	Medium-sized loop-type liquid metal fast breeder reactor
Loviisa	Finland	Loviisa PWR
Müh	Switzerland	Mühleberg BWR
P1100	Japan	Standardized 1100 MWe class PWR
P1300	France	Standardized 1300 MWe Framatome 3-loop PWR
P900	France	Standardized 900 MWe Framatome 4-loop PWR
Pick	Canada	Pickering A CANDU
PIUS	Italy	PIUS
SBWR	Italy	General Electric SBWR
Siz	UK	Sizewell PWR
TVO	Finland	TVO (Olkituoto) BWR
WS2/3/4	Korea (R.O.K.)	Wolsong Units 2/3/4 Pressurized Heavy Water Reactor (CANDU 6)

Table 6-3. Some characteristics of surveyed PSAs and HRAs.

Task Member	B	Canada	CH	CH	D	D	F	F
BWR			B					
PWR	P			P		P	P	P
Other		CANDU			GCR			
PSA Charact. ¹	1+ / I A	3 / I FP	2 / I,E FP	2 / I,E FP	3 / I FP	1 / I FP	1 / I A	1 / I A
HRA Methods Used For Cat. C Operator Actions	THERP amended by French-specific data	THERP expert judgement	FLIM	FLIM	THERP extended with features of HCR, SLIM, HEART, APJ	SHARP both ASEP and THERP	SHARP simulator data simulator-based TRCs ASEP	SHARP simulator data simulator-based TRCs ASEP

Task Member	I	I	I	J	J	J	NL	NL
BWR	SBWR			B			B	
PWR		AP600			P			P
Other			PIUS			LMFBR		
PSA Charact. ¹	3 / I,E A	3 / I,E A	1 / I,E A	2 / I FP	2 / I FP	1 / I,E + 2 / I FP	3 / I,E A	3 / I,E A
HRA Methods Used For Cat. C Operator Actions	SHARP1 ORE/HCR THERP	SHARP1 THERP	SHARP1 THERP	THERP OAT/TRC expert judgement	THERP OAT/TRC expert judgement	SHARP OAT/TRC THERP	OAT/TRC and THERP	SHARP ORE/HCR and THERP

Task Member	ROK	SF	SF	SP	UK			
BWR		B						
PWR			P	P	P			
Other	CANDU 6							
PSA Charact. ¹	2 / I,E FP (in progress)	1 / I,E A	1 / I,E FP	1 / I +FIRES FP	3 / I,E A			
HRA Methods Used For Cat. C Operator Actions	SHARP ASEP and THERP	SHARP - ²	simulator-based TRCs or THERP diagnosis curves expert judgement	SHARP HCR and THERP	HEART / THERP			

¹ PSA Level / I = Internal Events; E = External Events
Modes FP = full power; A = all modes

² SF- Probabilities from THERP diagnosis curves used as priors in a Bayesian update with simulator data or plant-specific judgement factors.

For reference, the methods commonly used in the U.S. IPEs were [NUREG-1560, 1996]:

- SLIM implementations (mostly PLG SLIM implementations)
- the EPRI cause-based decisions trees (“decision trees”, cf. Section 8.3)
- HCR and HCR/ORE
- a SHARP/HCR combination described by Dougherty and Fragola
- THERP or ASEP
- the Individual Plant Examination Partnership (IPEP) methodology (a modified THERP)

6.3 Critical Operator Actions

6.3.1 Overview

Appendix D lists the actions identified as important contributors in each of the surveyed PSAs. In general, the actions have been identified on the basis of the Fussell-Veseley importance measure, the percentage contribution to core damage frequency from the sequence(s) involving the operator action (the failure of the operator to perform the action). In a few cases, other importance measures, e.g. risk achievement worth, have been provided and are included in these summaries.

It should be noted that throughout this report, the term “operator action” is used to refer both to an action in a specific scenario and sequence as well as to a class of actions, that is, to similar actions in diverse situations. The intended usage is indicated where ambiguous.

The Fussell-Veseley importance indicates the contribution of the operator action (or class of action) among all contributors, that is, both hardware and operator actions. In PSAs and consequently, in the survey responses, important actions refer to the most important actions both among all contributors as well as among all actions. An action identified as important in the survey response is generally among the top 10 most important actions but may or may not be among the most important contributors overall when both hardware and human actions are considered.

On the basis of the important actions identified in the surveys, lists of important actions that are common to many PSAs were prepared. The following section shows these lists for BWR and PWR reactor types.

6.3.2 Important Actions in Common

Tables 6-4 and 6-5 list the important actions found in common for BWR and PWR reactor types. For BWRs, these are:

Operator action	Identifier
Manual Depressurisation	MD
Initiate Standby Liquid Control, in Anticipated Transient Without Scram (ATWS) scenarios	SLCS
Initiate Suppression Pool / Torus Cooling	SP/TC
Recover Residual Heat Removal (by restarting or starting alternate means of decay heat removal)	RHR

These operator actions are required in many different kinds of sequences. As will be the case also for the PWR common actions, the sequences in which the common important actions appear will frequently differ from PSA to PSA. As a result, caution is needed in any discussion of similarities or generic findings.

The two-part structure of the survey addresses the need for more detailed information. At the same time, it limited the number of actions for which this information needed to be provided.

For PWRs, the common important actions are:

Operator action	Identifier
Initiation of feed and bleed cooling.	FB
Response to Steam Generator Tube Rupture (SGTR)*	SGTR
Alignment of the emergency cooling systems to recirculation mode	RC
Recover Residual Heat Removal	RHR

* In most PSA models, the response to SGTR includes both the identification of a SGTR event as well as isolation of the affected steam generator.

Although the important actions from the PSAs for the SBWR and AP600 are included in these tables, the significant differences in these designs from more conventional light water reactors mean that the similarities in the actions are quite superficial. As a result, they are generally not discussed in this report.

Table 6-4. Important actions in common (BWR).

	MD Manual Depressurisation	SLCS Initiate Standby Liquid Control (in ATWS)	SP/TC Suppression Pool / Torus Cooling	RHR Recovery (i.e. restart or start alternate means)
B1100 (J)	x	x		x
DOD (NL)	x	x		
MÜH (CH)	x	x	x	
SBWR* (I)	x			
TVO (SF)	x	x		x

* Note the SBWR is an "advanced" BWR design.

Table 6-5. Important actions in common (PWR).

	FB Feed and Bleed	SGTR	RC Alignment for Recirculation	RHR Loss of RHR Shutdown
ALM (ES)	x	x	x	(n/a*)
AP600** (I)				(n/a*)
BEZ (CH)	x	x	x	(n/a*)
BOR (NL)	x		x	x
Doel (B)	x		x	x
DRS (D)	x			(n/a*)
LOV (SF)				(n/a*)
P1100 (J)	x	x		(n/a* SD study under way)
P1300 (F)	x			x
P900 (F)	x	x		x
SIZ (UK)	x			x

* These studies do not address low power and shutdown (LP&SD) modes of operation.

** None of these actions are identified as important in this study.

In general, these same actions were found to be important in the results of the U.S. IPEs [NUREG-1560, 1996].

“The only actions important in more than 50% of the BWR IPEs are manual depressurization, containment ventings, initiation of standby liquid control (SLC), and system alignment for decay heat removal. In PWRs, only switchover to recirculation, feed-and-bleed, and the actions associated with cooldown and depressurization are important in more than 50% of the IPEs.” [emphasis added]

6.4 Detailed Treatments

There is no intention to directly compare the HEP estimates since the plants represent a variety of designs and differences exist with respect to the degree of automation, procedures and context of the actions. However, some factors that drive the numerical values can be identified.

Additional, more detailed information was requested on the treatment of the actions that were found to be important in many of the PSA surveys. To broaden the comparison, detailed treatment descriptions of these actions were also provided for some PSAs in which these actions were not important.

Sections 6.4.2 and 6.4.3 discuss the treatments of these actions, for BWRs and PWRs respectively. Each of these sections includes tables summarising some key characteristics of the actions and results. Some general information about the summary tables is presented next.

6.4.1 *About the summaries of detailed treatments*

For each PSA, the importance of the action and its treatment is summarised by including

Scenario and sequence

- initiating event and sequence(s) in which the action appears

The sequence in which the action appears is described by listing the main failed systems.

Importance of action and sequence

- Fussell-Vesely and risk achievement worth importance measures

The Fussell-Vesely (F-V) importance measure is the fraction of core damage from sequences containing the operator action (or system).

The risk achievement worth (R. Ach. W. or R.A.W.) is the ratio

$$\frac{\text{CDF(with the action or system guaranteed failed, e.g. P}_{\text{failure}} \text{ or HEP} = 1)}{\text{CDF nominal}}$$

The R.A.W. is always larger than 1.

- sequence frequency

The sequence frequency is the cut set frequency. For actions that appear in multiple sequences, the sequence frequency is actually the sum of the frequencies of the cut sets. In many cases, the reported value has been calculated from the given Fussell-Veseley importance.

Characteristics of actions

- time available and median response time

The median response time is generally used only in quantification methods based on HCR and related methods. The median response time is in the majority of cases an estimate, i.e. an average of estimates of the required time to perform the action.

- important PSFs

Important PSFs, in this context, refer to those PSFs that are driving the HEP value. (The HEP is most sensitive to assumptions about these PSFs.) A full description of the PSFs considered and their rating is provided in the individual detailed responses (by PSAs).

Quantification

- the diagnosis and execution components of the HEP
- approach to recovery

The potential for recovering a failed operator action can be analysed in different ways. An explicit approach treats recovery as a third phase (DDD, Execution, Recovery). Implicit approaches consider, for instance, whether there is enough time to repeat the performance of the action. Where possible, the recovery factor is indicated along with the action phase (DDD or execution) to which it applies.

- methods used to assess the action

The detailed treatment surveys submitted for each PSAs provide much more information that is not (could not) be presented in the tables.

6.4.2 Detailed treatments for BWRs

For BWRs, the detailed treatments surveyed:

- manual depressurization (Table 6-6, 5 studies)
- the actuation of the standby liquid control system (SLIC) (Table 6-7, 2 studies + some information from a third study)

Manual Depressurization

Manual depressurization in BWRs is required in any sequence where high pressure systems for injecting coolant (feedwater) are not available. As a result, the scenarios involving this action are very diverse.

For manual depressurization the time windows used in the PSAs vary between 20 and 45 minutes; most (but not all) analyses consider stress associated with this action as high and consequently the most significant error producing condition among the PSFs. The estimated HEPs are predominantly of the order of 10^{-2} , although a much lower value (determined by the short median response time) applies to the Dutch Dodewaard plant.

As expected, the execution of manual depressurization does not contribute strongly to the failure of this action. The execution part has either been modeled using THERP, neglected or is considered to be implicitly covered by the integral evaluation (FLIM).

In the Finnish TVO study (base case), 41 minutes are available for the decision part (4 minutes within the 45-minute time window were allocated to the manual actions to execute depressurization). The 95th percentile curve of the ASEP nominal diagnosis model was used to establish a prior (4.9×10^{-3}) which was then updated using the results of simulator exercises to obtain the final, higher estimate of 1.3×10^{-2} . In the Japanese B1100 study, the time window is shorter but the full time available (i.e. 30 minutes) was allocated to the diagnosis period. The median curve of the ASEP nominal diagnosis model was used. This model as well as the absence of simulator data leads to a lower failure probability relative to the Finnish study.

In the Swiss Mühleberg study, the HEP values are determined by the selection of the calibration data. The primary results of the FLIM (SLIM) analysis consist of relative rankings of the modeled operator actions and qualitative information. For manual depressurization, the PSFs “Preceding and concurrent actions” and “Stress” are identified among the more important contributors to failure of this action. (Calibration, the conversion of the relative rankings expressed by the failure likelihood index into HEPs, is of course part of FLIM/SLIM. However, the HEPs depend completely on the selected calibration data.)

Some important characteristics that also influence the variability in HEPs for manual depressurization

- Is there an automatic depressurization system (ADS) active in the sequence being modeled? (Is this a back-up to the failure of automatic initiation?)
- Is the sequence of the type where the ADS would be automatically initiated but too late according to the PSA model (success criteria)?

SLIC Actuation

The actuation of the standby liquid control system (SLIC) is called for in Anticipated Transient Without Scram (ATWS) scenarios. Typically, this action is required early in scenarios, when the insertion of the control rods to scram the reactor fails.

For both the Swiss Mühleberg and the Japanese B1100 studies, the time available in these scenarios is rather short, 5 and 10 minutes respectively. The HEPs are relatively high; both are above 0.01.

Due to PSA model differences, however, the operator actions covered by these HEPs are significantly different. The B1100 study considers not only SLIC actuation but also inhibition of the Automatic Depressurization System (ADS) and control of the reactor water level (at a low level) as a single set of actions. In the Swiss study, on the other hand, the HEP models only the failure to actuate the SLIC. The remaining actions are considered separately.

In the Dutch Dodewaard study, SLIC actuation is not analyzed in detail in the study. All actions with time windows up to one hour are quantified with a TRC. Consequently, the HEP results directly from the 10-minute time window.

It is interesting to note that for the U.S. IPEs, the variability seen in 27 BWR IPEs, is of more than three orders of magnitude. The factors contributing to the variability were:

- some actions are manual initiations in cases of failure of automatic initiation, i.e. a recovery action (as was found for manual depressurization in this survey)
- early vs. late initiation of SLC, where late initiation has typically higher values in some cases because a dependency with the failure of early initiation is modeled.
- assumptions regarding the availability of the condenser, which change the time available

Table 6-6. Treatment of “manual depressurization” in the BWR PSAs surveyed.

Country-Study	CH - Müh	I - SBWR	J - B1100	NL - Dod	SF - TVO	
Initiating event	General transient (non-ATWS)	sLOCA	General transient (non-ATWS)	General transient (non-ATWS) and small LOCAs	Transient resulting in loss of all main feed pumps	
Sequence(s)	Feedwater and Reactor Core Isolation Cooling (HP Injection) failed	Failure of automatic start of HP Injection and depress. systems	Stuck open safety relief valve HP Core Spray failure	(Multiple sequences)	HP Injection failed	Stuck open safety relief valve HP Injection failed
Fussell-Vesely	1.2 E-1	3.0 E-1	4.3 E-2	7.7 E-3	2.4 E-2	2.9 E-3
R. Ach. W.	–	4.16	57	78.4	–	–
Seq.freq./yr	8.3 E-7	3.1 E-8	1.2 E-7	4.6 E-7	3.4 E-7	4.1 E-8
HEP	1.3 E-2	8.8 E-2	2.9 E-3	1.0 E-4	1.3 E-2	3.6 E-2
Time available	25 min.	35 min.	30 min.	40 min.	45 min.	20 min.
Median response time	n/a	15 min.	n/a	2 min.	n/a	n/a
Important PSFs	seven equally weighted PSFs ‘Preceding and concurrent actions’ and Stress rated worst	moderate dependence RO-SRO Stress level: ‘grave emergency’ (factor=5)	moderately high stress, good interface (for calculation of execution part)	“no burden” assumed	Stress related to decision-making burden Procedures and training good	
DDD	integral evaluation	8.7 E-2	2.7 E-3	1.0 E-4	1.3 E-2	3.6 E-2
Execution		9.7 E-4	2. E-4	neglected	considered negligible	
Recovery	implicit	credited (exec.)	credited (exec.)	implicit	implicit	implicit
Methods used to assess this action	FLIM	HCR/ORE for DDD THERP for execution part	THERP: -Nominal diagnosis median curve -Execution ¹	TRC	ASEP nominal updated with simulator data	

¹ Execution part of HEP includes recovery factors and hardware failures (HRA event tree).

Table 6-7. Treatment of “SLIC Actuation” in the BWR PSAs surveyed.

Country-Study	CH - Müh	J - B1100	NL - Dod
Initiating event	General transient (ATWS)	General transient (ATWS)	General transient (ATWS)
Sequence(s)	(Multiple sequences)		(Multiple sequences)
Fussell-Vesely	1.0 E-2	1.4 E-1	2.4 E-2
R. Ach. W.	–	–	1.77
Seq. freq. / yr	7.1 E-8	1.1 E-7	1 E-7
HEP (mean)	1.6 E-2	2.7 E-1	2. E-2
Time available	5 min.	10 min.	10 min.
Median response time	n/a	n/a	n/a
Important PSFs	adequacy of time, preceding and concurrent actions, training and experience, and stress	moderately high stress, good interface (for calculation of execution part)	n/a
DDD	integral	2.7 E-1	n/a
Execution	evaluation	3.7 E-3	n/a
Recovery	implicit	credited (exec.)	n/a
Methods used to assess this action	FLIM	THERP: -Nominal diagnosis -median curve -Execution	TRC

No other descriptions of detailed treatments available.

6.4.3 Detailed treatments for PWRs

For PWRs, the detailed treatments surveyed:

- feed and bleed (Table 6-8, Table 6-9, 9 studies)
- alignment for recirculation (Table 6-10, 4 studies)
- loss of Residual Heat Removal (RHR) (Table 6-11, 3 studies)

Feed and Bleed

The term “feed and bleed” refers to a mode of cooling using either primary coolant or feedwater. In primary feed and bleed, coolant is added and bled from the Reactor Coolant System (RCS). For example, emergency core cooling systems (ECCS) can be used to add coolant; steam is then released through the pilot-operated relief valves (PORVs).

In secondary feed and bleed, feedwater is added, e.g. to the steam generator, and (steam) released through the steam generator relief valves; this is also known as steam generator feed and bleed.

For “feed and bleed” (F&B), the time windows range from 20 to 75 minutes, with 5 minutes in the Spanish study as an outlier.

Analyses of the feed and bleed mode of cooling suggest that in some cases there may be some reluctance from the operators to initiate this mode of cooling for two potential reasons. First, primary feed and bleed releases coolant into the containment atmosphere, with associated costs for clean-up. Secondly, the release of coolant runs counter to the usual objective of maintaining sufficient coolant. As a result of potential reluctance, the operators may attempt to recover other systems before initiating feed and bleed or otherwise delay initiation of this mode of cooling.

The possible reluctance to initiate feed and bleed, is considered explicitly in a number of the analyses, although in different ways: by assessing a high level of stress with this operator action, by using curves for “difficult” diagnoses, or directly as a performance shaping factor. In the Swiss studies, this is reflected in the PSF “decision and diagnosis”. The French PSAs cite “possible reluctance” and the UK study “objectives conflict”. High levels of stress are also generally noted as affecting the performance of “F&B”.

Feed and bleed scenarios illustrate the difference between

- the operators properly understanding the situation and how to respond to it, and
- the operators performing the required action

Potential sources of reluctance are addressed as parts of safety culture and organisational factors. Furthermore, it is worthwhile to note that it is difficult to observe such reluctance in simulator exercises, which is one aspect of the “simulator” effect.

The probability estimates for the failure to establish “F&B” cooling range from 10^{-3} to 10^{-1} . This range reflects the use of different methods as well as the differences in the PSA event sequences in which these actions occur. For the detailed treatments from three studies, the sequence is a “general” transient. In

contrast, the action occurs in steam line break sequences in the Swiss study. The sequences in the French studies are Loss of Main Feed sequences with the failure of Auxiliary Feed. A steam generator tube rupture (SGTR) sequence is treated in the Japanese submission. Finally, the sequence from the Dutch study is a flood/fire during power operation that includes a LOCA through a Pilot Operated Relief Valve (PORV). The differences in these sequences underscore the need for caution when interpreting the discrepancies between the estimates for a single “common important action”.

For the purpose of contrasting the treatments of this action, the HEP is decomposed into three parts, 1) detection-diagnosis-decision (DDD), 2) execution, and 3) recovery. The approach for combining these parts differs from study to study. For instance, in the Japanese analysis, the DDD and execution components are combined, in the expected way, to result in a median HEP for “feed and bleed” of 1.8×10^{-3} . To take into account the failure of the previous action, the final HEP is obtained by taking the upper bound of the distribution (whose median is 1.8×10^{-3}), resulting in the estimate of 1.4×10^{-2} . In the French analyses, two types of recovery mechanisms are considered. The quantification of the execution part explicitly considers the recovery of the execution failure (mainly omission) by the operating crew. This recovery model considers a) whether an alarm is activated and whether a significant parameter is affected as a result of the omission, b) whether the procedure explicitly requires checking, and c) time available. The recovery shown in the table for the French entries refers to the separate analysis of intervention of the “safety engineer”. This model accounts for both the probability of the presence and the success in diagnosis (recovery) of the engineer.

Alignment for Recirculation

The continued operation of either low pressure or high pressure systems for injecting coolant depends on the availability of water to inject. The alignment for recirculation is required when the storage tanks of emergency injection water, from which injection systems initially rely on, and alternative sources, such as the refueling water storage tank, are depleted. The action is to realign the suction lines of the active injection systems to the containment sump, in order to recirculate the water that has been injected but has, for example, spilled back out of a break.

The time available for this action may be on the order of minutes, in large breaks where the reactor coolant system is quickly depressurized and the injection flow rates are consequently large, or of tens of minutes, usually in high pressure scenarios (smaller breaks).

The need for recirculation is normally covered in training. Some factors that drive failure include a high level of stress due to the severity of the plant condition and the need to manage concurrent actions and concerns. This action may take place both in scenarios where an accident evolution, while severe, is occurring according to training and expectations, as well as others where the need to assure a continued supply of injection water is only one of many concerns.

Consequently, the HEPs for this action ranged from 5×10^{-4} to 5×10^{-1} in the surveyed studies.

The U.S. IPE results also show that in addition to HRA assumptions, the scenario characteristics, particularly whether the switchover is at high pressure (small LOCAs) or low pressure (large LOCAs), are important factors driving the variability of the HEPs for this action. The high pressure scenarios are assumed to typically have more time available. In addition, the “difficulty and complexity of the general process to accomplish the switchover apparently varies significantly across plants.”

Loss of Residual Heat Removal

The operator response to Loss of Residual Heat Removal (RHR) during shutdown were described in three of the surveyed studies. All three actions occur in scenarios in which the plant is in the so-called midloop condition -- the water inventory in the primary is at its lowest level. Loss of RHR may then occur if level is not controlled or if there is a leakage, for instance due to an inadvertent letdown path. The low water level may then lead to cavitation and then tripping of the RHR pumps. In general, a main element of the response to this scenario is to provide make-up of the water level.

Although the details of the scenarios may vary, it can be seen that the appropriate response to this action is quite significant. In the surveyed studies, the failure of this action accounts for 11 to 17% of the core damage frequency.

In the French P900 and P1300 studies, the action refers to both the recognition of the condition as well as the execution of the actions to provide make-up water. The probability is driven mainly by the time available and obtained from the EDF diagnosis curves (TRCs).

In the UK Sizewell study, the surveyed action refers only to the diagnosis of the condition. It is modeled with the THERP annunciator response model with credit for recovery by the supervisor. The response or execution part is modeled separately. The HEPs for the actions modeling the execution part (not shown in this table), which apply to different plant conditions, range from 5E-5 to 2.6E-3.

Table 6-8. Treatment of “feed and bleed” in the PWR PSAs surveyed. (p. 1 of 2)

Study	B - Doel	CH - Bez	F - P900	F - P1300	J - P1100
Initiating event	General transient	General transient Steam line breaks in-/outside containment	Loss of main feed (MFWS)	Loss of main feed (MFWS)	SGTR
Sequence(s)	Reactor trip and total loss of SG Feedwater (AFW, EFW, and MFW)	Reactor trip with failure to trip of at least one main turbine and failure of MSIVs to close (“Severe, rapid overcooling”)	Failure of aux. feed (AFWS)	Failure of aux. feed (AFWS)	Successful trip, high pressure injection, aux. feed Failure to isolate faulted SG
Fussell-Vesely	1.8E-2	2.6 E-2	< 6.6 E-4	1.0 E-2	9 E-2
Seq. freq. / yr	4.2E-7	2.9 E-7 (mult. seqs)	3.3 E-8	1.5 E-7	3.9 E-7
HEP (mean)	1.8E-2	1.6 E-1	1.0 E-3	5.6 E-3	1.4 E-2
Time available	20 min.	20 min.	75 min.	60 min.	30 min.
Median response time	n/a	n/a	n/a	n/a	n/a
Important PSFs	local action (unfavorable)	Stress Decision and Diagnosis	Possible reluctance (“difficult” diagnosis curve used)		High stress Failure of previous action
DDD	1.8E-3	integral	5.0 E-3	7 E-2	1.0 E-3
Execution	1.6E-2	evaluation	1.8 E-3	considered negligible	9.2 E-4
Recovery	credited	implicit	1.5 E-1 (non-recovery)	7.9 E-2 (non-recovery)	credited (for exec. part)
Methods used to assess this action	EDF (French) methods as in EPS-1300 (1990) (cf. <i>F - P1300</i>)	FLIM	EDF diagnosis curves (TRC) Simulator-based generic values		THERP: - nominal diagnosis curve (median) - Execution

Table 6-9. Treatment of “feed and bleed” in the PWR PSAs surveyed. (p. 2 of 2)

Study	NL - Bor	SF-Lov	SF-Lov	SP - Alm	UK - Siz
Initiating event	Flood or fire in an instrument room during power operation	Loss of Instrument Room Ventilation (LIRV)	Any transient (General transient)	General transient	General transient (inadvertent reactor trip)
Sequence(s)	LOCA through PORV Failure of EECS to start automatically Failure of operator to initiate secondary cooldown within 30 min.	Reactor trip and RCP seal coolant bleed-off isolation both succeed	Feedwater and Emergency Feedwater both failed (Total Loss of Feedwater)	Failure of secondary circuit heat removal	Failure to relieve secondary side overpressure OR Failure of short-term decay heat removal via main FWS and of aux. feed
Fussell-Vesely	8.4 E-2	5.3 E-2	5.1 E-2	< 3.0 E-2	1.5 E-1
Seq. freq. / yr	4.9 E-6	1.6 E-6	1.5 E-6 (mult. seqs)	5.1 E-7	not available
HEP (mean)	3.0 E-1	1.4 E-2	9.4 E-4	8.4 E-3 (median)	2.2 E-2
Time available	70 min.	n/a	n/a	5 min.	60 min.
Median response time	40 min.	n/a	n/a	50 s	n/a
Important PSFs	Stress	Level of simulator training	Level of simulator training	Stress: grave emergency	Unfamiliarity Objectives conflict
DDD	2.9 E-1	commission 1.4 E-1	diagnosis 1.4 E-4 commission 8.0 E-3	7.0 E-3	1.9 E-2
Execution	9.0 E-3			1.45 E-3	3.0 E-2
Recovery	implicit	0.9 included	0.9 included	implicit	not credited
Methods used to assess this action	HCR/ORE w/ simulator-based “decision trees” Generic values	expert judgement (IVO structured approach)	TRC for diagnosis + IVO structured exp. judg. approach	HCR THERP: - Execution	HEART

Table 6-10. Treatment of “Alignment for Recirculation” in the PWR PSAs surveyed.

Study	B - Doel	CH - Bez	NL - Bor	SP - Alm
Initiating event	Large Break LOCA	SGTR or PTS	Small break LOCA	Small LOCA
Sequence(s)	Reactor trip	SGTR sequences with core melt. Other sequences leading to pressurized thermal shock (PTS). Long-term debris bed cooling and containment heat removal.	Failure of automatic start signal for sump recirculation	Reactor trip, HPSI, Secondary circuit heat removal, RCS depressurization (all successful)
Fussell-Vesely	6.2E-2	7.7 E-2	4.5 E-3	6.8 E-2
R.A.W.				83.8
Seq. freq. / yr	1.4 E-6	6.6 E-7 (mult. seqs)	2.5E-7	7.1 E-6
HEP (mean)	1.2E-3	5.2 E-1	1.3 E-2	5.1 E-4 (median)
Time available	2 min.	25 min.	25 min.	n/a
Median response time	n/a	n/a	3 min.	n/a
Important PSFs		Stress Decision and Diagnosis		n/a
DDD	Neglected	integral	1 E-2	n/a
Execution	1.2 E-3	evaluation	3.0 E-3	n/a
Recovery	credited	implicit	implicit	n/a
Methods used to assess this action	EDF (French) methods as in EPS-1300 (1990) (cf. <i>F - P1300</i>)	FLIM	HCR/ORE w/ simulator-based “decision trees” Generic values	HCR THERP: - Execution

F - P900, F - P1300 : treated “Loss of RHR”

J - P1100, SF - Lov, UK - Siz : this action not covered

Table 6-11. Treatment of “Loss of RHR” in the PWR PSAs surveyed.

Study	F - P900	F - P1300	UK - Siz
Initiating event	small LOCA during midloop	excessive draining during midloop	Loss of RHRS during midloop
Sequence(s)	Loss of RHR	Loss of RHR	
Fussell-Vesely	1.1 E-1	1.6 E-1	1.7 E-1
R.A.W.			342
Seq. freq. / yr	5.5 E-6	2.4 E-6	
HEP (mean)	4 E-2	1.1 E-3	5.0 E-5
Time available	30 min. for diagnosis 20 min. for exec.	60 min.	not available
Median response time	n/a	n/a	n/a
Important PSFs	Action is executed locally.	Training is rated good.	
DDD	1 E-2	5 E-3	1 E-4
Execution	4 E-2	2 E-3	modeled separately
Recovery	considered but not credited due to insufficient time; no procedure for safety engineer	recovery of execution failure credited Pnr=0.1; safety engineer also credited Pnr=0.16	credited, Precovery=0.5
Methods used to assess this action	EDF diagnosis curves (TRC) Simulator-based generic values		THERP

6.5 Comments on “Own” Methods

This section presents the comments on “own” methods, that is, on the methods *and their application* in each of the studies.

The comments on “own” methods represent the views of the individual contributors. They are provided here as submitted, i.e. they have not been edited.

6.5.1 Canada - Pick

The HRA methods developed and used in Canada are described in Section 0, which discusses the evolution of the methodologies and their applications.

In the Pickering A Risk Assessment the technique used for preliminary quantification of Category C human errors considers three dimensions as important determinants of error probability, namely: task characteristics, quality of indications, and time availability. Task characteristics are considered to be one of 3 kinds, straightforward and familiar, of average complexity, or very complex. Indication of need for action may be either unambiguous, require interpretation, unclear, or non-existent (4 possibilities). Time available for diagnosis and execution may be either unrestricted, greater than required, about equal to required, or less than required (also 4 possibilities). A table with 48 cells, covering the various possibilities above has been developed to permit the selection of HEPs once the task attributes are established. Guidance is provided in the methodology to determine how a given task should be tested against the above attributes.

Dependency between post-accident human interactions (HIs) is accounted for at the sequence cutset level. Each post-accident HI is classified as either being event sequence-specific, or related to the generic functions of reactor shutdown, heat transport system cooldown, and containment isolation on which operators are regularly trained. HIs of the same kind in an event sequence are considered to be completely dependent, while HIs of different kinds are taken to be independent unless they are postulated to occur close to each other in time, in which case only the generic type is retained as it is assumed the operator will give priority to protecting basic safety functions over correcting specific equipment failures.

Credit for recovery before an accident sequence leads to core damage or a release of radioactivity to the public is also incorporated at the integrated sequence cutset level. Recovery actions considered are either recovery from the initiating event, or correction of equipment failures in sequences where at least two hours are available between failures of all mitigating systems and core damage. Among the former, recovery is applied on only those initiating events for which the possibility of recovery is suggested by operating experience. For obvious reasons, recovery is not credited for pipe failure events except if the leak rate is very small. Mitigating system equipment failures are considered correctable only if the failed component is located outside containment and relatively easy to repair, e.g., restoring a power supply to a control component.

6.5.2 Czech Republic

The general procedure SHARP for human reliability analysis in frame of PSA recommended by IAEA was used in the NPP Dukovany and NPP Temelín PSA studies. Although the level of detail of several steps of SHARP was slightly reduced in comparison with the original version and some up-to-date methods of quantification not mentioned in SHARP were used, the general framework of SHARP was followed.

Within SHARP, the screening part of analysis is generally supposed to be of high importance, therefore screening was adequately addressed in both PSA studies. In the Dukovany PSA, the approximate method TESEO was used for screening. In the Temelín PSA, the „screening part” of the method ASEP was used.

During screening, the procedure-driven post-accident operator actions proved to form almost the whole contribution to the risk of plant operation connected with human factor; that’s why this category of human interventions was analyzed in detail in the subsequent analysis. The method THERP was planned to be used for the detailed analysis in case of NPP Dukovany.

An EPRI method of decision trees was suggested by US specialists for human reliability analysis in the frame of the Temelín PSA. This method was proposed because it could address better the potential problems with cognitive activities of control room staff (i.e. information processing and evaluation), which are very important for human reliability. The method was modified to address the specific issues of control of NPP Temelín operation, which have not been studied in the original version of the method (computer system COMPRO significantly helping operators to diagnose the non-standard status as well as to achieve the planned goals). For the analysis of execution part of control room staff interventions, the modified method ASEP was used.

For the EPRI method of decision trees was found to be very flexible and useful in case of NPP Temelín PSA, it was decided to be used in case of NPP Dukovany PSA, as well. The method was modified to take into account the specific features of NPP Dukovany (or more general - WWER-440) operation. A high stress was put upon the control room ergonomics, the level of experience of the operators and the quality of emergency procedures. Similarly to NPP Temelín PSA, the method ASEP was used for analysis of the manipulative part of human interventions.

6.5.3 Finland - TVO

The method used in HRA in the rev. 1 of the TVO PSA is developed in VTT by Pekka Pyy and a description of the method is presented in the Task questionnaire response (rev. 1, 17 May 1995). The method is also shortly described in a PSAM-III paper by Pekka Pyy and Risto Himanen (Vol. 2, p. 882).

The method is transparent and it is easy to follow the different steps of the method. The method combines in a clear way the use of different kinds of evidence, i.e. results of simulator tests and expert judgement.

If results of simulator tests are not available, performance shaping factors (K-factors in this case) are used to modify the point estimates of Swain. In the use of K-factors similar theoretical problems exist as in all other PSF-based methods. Especially concerning the factor "stress" one can ask, whether it is independent of the other factors. No basis has been presented concerning the selection of the values that are used for these factors. The calibration of the results, when using the K-factors, is a problem. Therefore, we have urged the utility to take other sequences, analysed in HRA, into their simulator re-training courses and to compare results achieved using both methods (the method based on the beta-distribution and the method based on K-factors). The utility is considering this issue at the moment.

Concerning the application of the method a positive aspect is that always a qualitative description and a basis for the assessment of each K-factor was provided. In the assessments more than one expert could be used. Some remarks concerning the detailed application of the method can be presented.

The use of Swain's nominal curve as a basis for the a priori can be questioned, because in some comparisons it has been shown that it often provides the most optimistic results. However, the 95% curve

was used for available times over 30 min. The experience from the use of K-factors was that they usually increase the human error probability.

As a positive point it can be mentioned that a correction factor is included in the method to take into account, if necessary, the difference between behaviour at simulator and at plant.

6.5.4 France - P900/P1300

Origin Of The Method

The method used in the French PSAs was mainly developed by EDF, in concert with IPSN. This method relies as far as possible on French experimental data when available (simulators - real experience), completed by generic data (Swain data).

Strength Of The Method

The important experimental basis was very useful for the credibility of the results.

The realism of the study was a major preoccupation. Therefore EDF gathered as much information as possible, directly at the plant or indirectly via a representative in the plant (who performed surveys and gave a lot of information). We also used 204 simulator tests performed with 78 teams of EDF operators.

The method was easy to use, with a rather limited number of parameters needing judgement. It was considered that the ranking of human factor contributions is more important than the absolute values.

Improvements Needed

After some years of PSA applications, improvements appear as particularly necessary :

- pre-accident errors (category A) : experience feedback indicates that dependencies between pre-accident errors are insufficiently analysed and probably under estimated ;
- accidental errors (category C).

Although the HRA quantification based mainly on experimental data has the interest of realism, however this approach has also limitations, especially for data collected by simulator tests.

In addition to the problem of transferability of simulator data to real situations, it can be noted also that it is not possible to collect data for all the situations quantified in the PSA, and it is then necessary to « extrapolate » the data to other situations, with help of judgement. Some examples of difficult problems which need further studies are the following :

- the case of long delays (more than 30 minutes) for diagnostic or for recovery ;
- dependencies between actions (between different actions performed by the same operators or between different operators) ;
- effect of actions carried out locally ;

- effect of parameters other than time (training, organisation, man-machine interface). Due to these problems it is very difficult to assess a priori the benefit due to improvements of these parameters.

6.5.5 *Germany*

In this section, limitations and constraints of HRA methods are discussed. The criticism relies on the limitations of the HRA methods as they were observed in the application of the methods to Human Reliability Assessment (HRA) problems in Germany.

Though the following is focusing on practical limitations of HRA methods, this does not imply that the HRA-methods are of low value or useless for HRA. Currently no other HRA methods are available that are of equal predictive value and are as established as these methods for the assessment of human operations.

In Germany, a limited set of HRA Methods is currently applied for regulatory purposes of Human Reliability Assessment. In most cases THERP is used for HRA /1/. In cases where information is lacking or where conservative assessments have to be made (i.e., when screening values are sufficient), ASEP is used too (e.g., /2/). See /4/ and /16/ for HCR, /5/ for THERP and /6/ for ASEP. Recently also HCR was introduced for the assessment of accident management situations /3/. For discussion of the methods see also /9, 14, 15, 11, 12, 18/.

The assessment procedure is widely independent from the HRA-method and closely related to the SHARP procedure /7/. Before quantification, the results of the qualitative study (plant and system behaviour, study of procedures, visits of plants, questionnaires to operators) are incorporated into a model of the operators' behaviour. This model depicts the expected behaviour of the operators and the system. It contains all important information for further assessment: for instance, tasks and possible task conflicts, ergonomic constraints, and organisational aspects.

6.5.5.1 *Detailed view on the limitations of the methods*

The problems of HRA methods may be discussed reasonable by looking at different ways of Quantification. For this reason, HRA methods may be distinguished into two major groups /8, 9/: decompositional methods (for instance, THERP and ASEP) or holistic Methods (such as HCR).

Decompositional HRA methods decompose the event into different sub-elements (i.e., into a human error event tree). Based on the sub-elements they assess and quantify the event according to the accompanying PSF (performance shaping factors) that were found for the sub-elements. A clear advantage of this detailed assessment is that most effective improvement-measures may be found by looking at the most important unavailability (e.g.: the correct Diagnosis is the leading error contribution of the whole sequence).

Main effort of THERP is the identification of different error likely situations for different action-steps. The most obvious problem of a procedure like this is the high effort of defining the sub-elements that are of interest for system failure (especially for scenarios where the sequence of actions is very long). This procedure not only takes much time but also needs a high amount of analysis-knowledge and expert experience.

The focus on single errors in THERP leads to the problem that the view for the global situation or error context for the operator (like organisational factors) or more general influencing factors (like management for instance) may be lost because a systematic broader view to the entire sequence is missing. THERP does not provide a detailed description how to consider such global influences in task analysis. This makes an assessment of complex procedures difficult (e.g., steam generator tube rupture, feed and bleed).

Another less stated problem is the difficulty to assess the dependencies between the defined sub-elements. THERP for instance is only assuming the dependence between two succeeding tasks according to the dependence model or between persons by the recovery-model. Dependencies that cannot be modeled by these approaches can only be considered by engineering judgement. Since dependencies are more complex (e.g., making the same errors in different redundancies), this limited opportunity is too restricted /10/.

Another aspect is the consideration of recoveries. THERP assumes recoveries e.g. by alarms, additional personnel, announcements, procedures. Possible recoveries of an error by the person who made the error are not considered. In cases where recoveries of failed actions are becoming possible after some time, this restriction may lead to pessimistic results (THERP assumes complete dependence of failure in action i at time t and failure in action i at time $t+x$).

Therefore, THERP not only has problems in applications in complex scenarios due to high effort but also due to principal problems of assessing important global error likely situations and dependencies.

Another aspect of application is the often stated limitation of THERP to skill- and rule-based actions and the exclusion of knowledge-based actions. Looking at the quantification process of THERP, this limitation has to be re-considered and treated in more detail:

(1) Psychological findings indicate that nearly every human action is influenced by knowledge-based elements and that the distinction into these three levels of behaviour is more artificial than real /13/. THERP is indeed considering this smooth relation between the levels of behaviour and is not limiting itself to skill- and rule-based actions because it is assuming a log-normal distribution of the error probability. This means that the median human error probability is representing the assumption of THERP that operators are usually well-trained and are performing planned actions (i.e. skill- and rule-based actions). Beside model and data uncertainties, the uncertainty bounds describe deviations from this average-level of behaviour either in the direction of very high skill (lower bound) or knowledge-based influence (upper bound), see /14/ for this discussion.

(2) Independent from this cognitive strain is the situation where the action is to be performed. THERP only provides data for situations with regular and frequent planned actions but not for unplanned actions in unusual situations (cf. Table 5-10 of this report). Hence, following this psychological perspective consequently would lead to the following requirement: It is necessary to ask operators before quantification whether they are familiar with the situation or would perceive it as an unusual situation. It is not sufficient to conclude from the existence of a written procedure familiarity with the situation and hence deduce the level of behaviour the operators will use. Otherwise, asking operators about how to cope with an accident management (AM) measure may lead to the result that they may be able to perform it on a rule-based level if it is sufficiently trained. In those cases, AM measures may be assessed by THERP (cf. /3/).

Holistic HRA methods such as HCR are assessing the scenario as a whole. Here, the total unavailability is not calculated from the unavailability of sub-elements but by estimating an unavailability for the whole scenario. To get enough face validity these methods need a well-elaborated model of the action sequence that the operators have to perform during the accident (model of the operators' behaviour). Having elaborated this model, the holistic methods only need low effort for quantification. Holistic methods seem not to have restrictions and problems of assessing global influences in complex scenarios like THERP has.

The main effort for HCR is the development of the model of the operators' behaviour and not the preparation of the assessment since HCR uses only a few parameters for quantification. Because one has to perform a detailed qualitative analysis before making the assessment, holistic methods obviously do not result in lower effort for the overall assessment procedure.

Furthermore, HCR is limited in addressing improvement measures by the quantitative assessment. Except by trying to change the values for the model parameters, measures of improvement may only be derived from the model of behaviour and not from the quantification itself. Hence, quantification is restricted in providing additional information for improvements and in providing more decision aids concerning specific aspects of the entire task. Other practical limitations are:

- HCR gives no support for finding error likely situations like THERP does it for instance. Its major use is quantification. The other part of HRA (the qualitative analysis of potentials for improvements) is not included. However, SHARP is recommended.
- HCR is only of use for Type C actions (i.e., post initiators) that do not deteriorate the situation.
- To use HCR for complex scenarios, the leading error contribution must be the diagnosis part of the task. This means that the actions on the plant have to be negligible. If this is not the case, HCR cannot be used for assessment alone and other methods have to be used for assessing the actions /16/. Typical examples are very error likely actions or unexpected recoveries (e.g. by crisis management).
- Already often stated is the problem of the limited number of PSFs. Three parameters with five distinctions at maximum give no detailed view on PSFs of a complex scenario. Helps for applying the PSFs are not well developed and the origin of the classification is unclear. Also the interrelations and side-effects of PSFs are not considered. Remind that HCR uses the parameter K_1 for the experience level of the crew with the task, K_2 for the expected stress level, and K_3 for the quality of the control room design.
- Dependencies between actions or individuals are only considered globally and unspecifically. Effects of different organisational solutions cannot be predicted due to missing PSFs for that field of ergonomics.
- Beside the problem that the normalised time assumes an independence of operators' reliability from the absolute available time /15/, HCR is also only applicable for scenarios where the absolute time windows are not too long. Hannaman /16/ gives about one hour as an example for the maximum time for the diagnosis part of the scenario since this is supported by NUREG/CR-3010 /17/. This leads to problems if one tries to use HCR for

assessment of recoveries e.g. in low-power and shut down modes because here the time windows are usually longer.

6.5.5.2 Literature

- /1/ **DRS-B (1990)** Deutsche Risikostudie Phase B. BMFT (Hrsg.) TÜV Rheinland Verlag. Köln.
- /2/ **SWR1 (1993)** SWR-Sicherheitsanalyse Anschlußbericht Teil 1. GRS-102/1. ISBN 3-923875-52-5. GRS. Köln.
- /3/ **SWR2 (1995):** Frey, W., Gänßmantel, G., Heinsohn, H. Hofer, E., Holtschmidt, H., Kersting, E., Kreuser, A. v.Linden, J., Mayer, G., Piljugin, E., Pointner, W., Preischl, W., Sträter, O., Versteegen, C., Bongartz, R., Reer, B. & Ullwer, W. (1995) SWR-Sicherheitsanalyse, Phase II. Abschlußbericht, Band 1: Untersuchungen von Ergebnissen aus dem Leistungsbetrieb. GRS. Köln.
- /4/ **Hannaman, G. W. & Spurgin, A. J. (1984a)** Human Cognitive Reliability Model for PRA Analysis. NUS-4531. NUS-Corp. San Diego.
- /5/ **Swain, A. D. & Guttman, H. E. (1983)** Handbook of Human Reliability Analysis with emphasis on nuclear power plant applications. Sandia National Laboratories, NUREG/CR-1278. Washington DC.
- /6/ **Swain, A. D. (1987)** Accident Sequence Evaluation Program on Human Reliability Analysis Procedure. NUREG / CR-4772. NUREG.
- /7/ **Hannaman, G. W. & Spurgin, A. J. (1984b)** Systematic Human action Reliability Procedure (SHARP). EPRI NP-3583. EPRI. Palo Alto. California.
- /8/ **Zimolong, B. (1991)** Empirical Evaluation of THERP, SLIM and Ranking to Estimate HEPs. Reliability Engineering and System Safety. Vol.35-1. Elsevier. p. 1.
- /9/ **Reer, B., Sträter, O. & Mertens, J. (1996)** Evaluation of Human Reliability Analysis Methods Addressing Cognitive Error Modelling and Quantification. Berichte des Forschungszentrums Jülich; 3222, KFA, Jül-3222, Jülich (D)
- /10/ **Reer, B., Bongartz, R. & Ullwer, W. (1995)** Zuverlässigkeitsanalyse von Personalhandlungen bei Störungen des Leistungsbetriebes eines Siedewasserreaktors - insbesondere von Notfallmaßnahmen. KfA-IST-IB-7/95. KFA. Jülich.
- /11/ **Sträter, O. (1996b)** A Method for Human Reliability Data Collection and Assessment. ESREL '96 / PSAM-III. Crete, Greece, June 24-25, 1996.
- /12/ **Sträter, O. (1996c)** Assessment Of Cognitive Errors And Organisational Aspects Based On Evaluation Of Plant Experience. PSA'96 Conference in Utah (USA) September 29 - October 3, 1996.
- /13/ **Wickens, C. D. (1984)** Engineering Psychology and Human Performance. C. E. Merrill Publishing Company, A Bell & Howell Company. Columbus, Toronto.

- /14/ **Sträter, O. (1996a)** Beurteilung der menschlichen Zuverlässigkeit auf der Basis von Betriebserfahrung. Dissertation. Eingereicht am 21.5.1996. Technische Universität München. München.
- /15/ **Sträter, O., Preischl, W. & Berning, A. (1994)** Untersuchungen zu speziellen Methodenfragen im Bereich Personalhandlungen. GRS Abschlußbericht. Nr. 2151. GRS. Köln.
- /16/ **Hannaman, G. W. (1987)** Use of Human Reliability Analysis for PSAs and Plant Application. Draft contribution to a manual for Probabilistic Safety Analysis and its Application in Safety Decisions. International Atomic Agency, Division of Nuclear Safety. Vienna, Austria. Revised Draft September 1987.
- /17/ **Hall, H. E., Fragola, J. & Wreathall, J. (1982)** Post-Event Human Decision Errors: Operator Action Tree/Time Reliability Correlation. NUREG/CR-3010. NUREG. Washington/DC.
- /18/ **Mosneron-Dupin, F., Reer, B., Heslinga, G., Sträter, O., Gerdes, V., Saliou, G. & Ullwer, W. (in press)** For more Human-Centered Models in Human Reliability Analysis: Some Trends Based on Case Studies. submitted for publication in Reliability Engineering and System Safety. Elsevier.

6.5.6 *Italy*

Criteria in the choice of HRA methods

In the PSAs performed, methods from the literature have been selected depending on the plant for which the analysis has to be performed. The studies are performed as much as possible accounting for “plant-specific” factors.

Thus, the tasks of the analysis can be different and/or performed in different ways if the plant is operating, is under construction or is in the conceptual stage. As an example for Caorso, that at the time of the analysis was operating with about 10 years of operation, we performed an extensive task of operator interviews in order to understand the response to the different situations, timing of the actions, the real use of the written procedure, etc. The HCR approach supplemented by THERP was followed.

In Alto Lazio, that was under construction at the time of the analysis, THERP was followed but a very large use of screening, conservative values for the human errors was performed: these values were refined only if they were found to be important from the point of view of the contribution to the overall core damage frequency or to the frequency of high release to the environment.

For the new passive reactor developed by GE and Westinghouse (SBWR and AP600), the EPRI approach (HCR/ORE) was followed, but also in this case, because the plants were in the conceptual stage and one of the main goal was the demonstration of the independence by the operator intervention (quantification of the core damage frequency has been performed also for the case in which no operator intervention at all was considered) a large use of screening and conservative values were used.

Practical problems implementing methods

Some problems have been found in implementing the HCR/ORE approach for the new passive plants because many actions were not simulated in the ORE experiments. So, in addition to conservatism,

engineering judgement had to be used for some actions for the estimation of the average time for performing the action and of other necessary parameters. Also engineering judgement had to be used to estimate the probability p_1 in the cognitive part of the actions (this probability has not been any more considered in the revised approach reported in attachment 2).

For the other methods, no practical problems have been found.

Own views on methods and their application

All methods, in principle are „good”, but they, as demonstrated by international benchmarks, can be applied in very different ways (for example, even THERP can give very different results if the operators are considered more or less dependent, if the verification of the action is modelled or not, etc.) and so, can give very different results. This different application depends on many factors not dependent on the method itself as for example the experience of the analyst, the effort put in the HRA (percentage of the resources), the plant examined, the regulatory regime of the country in which the plant is located, the quality of the man-machine interface, the degree of automatization of the safety equipment, the organisation of the control room staff in terms of responsibility,, etc.

6.5.7 Japan - B1100/P1100

type 3:

- Dependencies between tasks are evaluated by engineering judgement, namely by using the upper uncertainty bound of HEP for the succeeding task when the preceding task fails, considering reduced acceptable times for the succeeding task. The preceding task is supposed to actions indicated in the event-based procedure and the succeeding task, in the symptom-based procedure, respectively. There are grounds for investigation in this area.

type 4:

- Inadvertent termination of mitigation systems by operators before termination criteria is not considered in the level-1 PSA. It should be appropriately evaluated if these importances are significant.

type 5 :

- Some repairs or restorations of failed components could be conducted as a matter of course. For example, when the failure of control circuit of isolation valve is identified, operators should open the local valve manually, of which actions may be significant risk reduction contributors. But these actions may be improvised, and the procedures, training, education etc. are ambiguous in actual plant. These kinds of operator actions are not estimated in the level-1 PSA.
- In Japan, some accident managements, such as alternate RPT and ARI, use of electric power of adjacent plant, alternate water injection into core, inforced pressure-proof vent are now being developed to prevent and to mitigate severe accident. In the course of selecting proper accident managements in individual plants, efficiencies of accident managements for risk reduction are useful information However, accident management actions are a kind of

dynamic task in the sense of substitution of the proceeding actions. we must establish proper human error methodology.

6.5.8 *Netherlands - Bor*

These remarks represent the opinion of M.F. Versteeg, of the Nuclear Safety Department (KFD), The Netherlands, and do not necessarily reflect the opinion of the regulatory body.

- The part of the PSA dealing with Human Interactions has an adequate level of detail and is of good quality. Moreover, the assessment of the Errors of Commission is even innovative and of high standards compared to the State of the Art. Therefore, in connection with the sensitivity analysis the results can be used for those kind of LPSA applications dealing with operator actions, such as: Review of Normal & Emergency Operating Procedures, backfitting (automation of some critical operator actions), Optimising Test & Maintenance, Optimising or modification of Operator Training Programmes, etc. Although, not too much emphasis should be put on the numerical outcomes as such.
- Compared to similar actions in other PSA studies, some operator actions tend to have lower values, especially those with relative very long time windows and short median crew response and manipulation times. These lower values are solely due to the use of the HCR-ORE model in this domain. Values for cognitive processing/procedural mistakes (P1) < 10⁻⁴ look very optimistic, even for situations with clear presentation, well practised by the operator and ample time for cognitive processing and recovery (see table 3). Biasing based on previous events is the accident sequence, group dynamics, etc., might be influences which act counter-productive for the cognitive processing. Despite these (too?) optimistic numerical values of some HEPs. the insights gained by the assessment of the operator actions remain valid.
- A weakness of the HCR-ORE (and almost any other HRA model) is the excessive dependency on expert judgement. Especially, answers to questions related to whether or not a particular operator action is stressful, what the length of median crew response times and manipulation times are, were derived by interviewing control room personnel and simulator trainers. Overconfidence might easily have slipped into these answers.
- A "weakness" in the study is the treatment of the Long Term Recovery Actions via repair actions, local actuation of components in case actuation from the control room is disturbed, or the use of alternative components and/or systems via cross-ties. The State of the Art in modeling these kind of actions is currently not so developed that dependencies of those actions on the sequence history and/or the associated environmental conditions can be included. Especially, this is the case with recovery of damaged components/systems resulting from area and external events where plant situations might become unclear to the operator. Several of these events are listed with screening values for the HEPs. The use of Markovian models might be a solution for this case.

6.5.9 *Switzerland - Muh/Bez*

In three of the four Swiss PSAs, the Success Likelihood Index Methodology (SLIM) is applied to the quantification of the Cat. C (dynamic) operator actions. The comments provided in this supplement refer

primarily to the application of SLIM in the Mühleberg PSA. The application of SLIM in the Beznau PSA is also addressed briefly below.

It should be noted that the implementations of SLIM by PLG Inc. differ significantly from the methodology as proposed originally by Embrey et al. (Embrey et al., 1984), which is referred to as SLIM-MAUD. The PLG SLIM implementation in the Mühleberg PSA is essentially the one documented by Chien et al. (1988) as well as in NUREG/CR-6144, the Low Power and Shutdown Study of Surry, Unit 1 (Chu et al., 1994). (PLG Inc. has also provided several SLIM versions.)

The key elements that differentiate PLG SLIM from SLIM-MAUD are:

- In PLG SLIM, a number of groups of operators (experts) are asked to judge the performance shaping factor (PSF) ratings and weights.
- The PSFs used to make the relative ratings are chosen by the analyst.
- PSF scaling guidance is provided to the experts in the form of anchored rating scales.
- PLG SLIM is defined in failure space, that is, failure likelihood is used instead of success likelihood.

Dang and Hirschberg (1995) discuss in more detail the differences between 'PLG SLIM' and SLIM-MAUD.

Some of the positive features of PLG SLIM are:

1. Clear documentation of the context of the operator action is emphasised, including both detailed descriptions of the specific groups of sequences in which the action takes place as well as the procedural support for the operator response. A template is used for these operator action descriptions, helping to ensure that the key factors are considered.
2. The input of plant operators is elicited in the assessment of the relative difficulty of operator actions. The potential tendency of operators towards optimistic judgement of failure probability is avoided because these are relative judgements.
3. In addition, the PSF scaling guidance should promote consistency and repeatability in the subjective judgements of the experts.

The key issue for all applications of SLIM is the calibration step, which allows the relative rankings of failure likelihood to be converted into failure probabilities. The calibrations in the Mühleberg study are based on probabilities from other PSAs, which raises questions of applicability and transferability.

A second, though less critical issue, is that the selected set of PSFs tends to focus on the decision aspects of the context, for instance, relative to control room man-machine interface issues. The desirability of such a focus would need to be assessed when SLIM is considered for other PSA studies.

Summary

The use of systematically elicited input from plant operators is a major advantage of the method. The relative difficulty judgements obtained are a useful result and moderate to some extent the critical issue of

calibration. Studies have shown that the multi-attribute decomposition of the difficulty into its component PSFs as done in PLG SLIM leads to consistent assessments among the experts. Concerning calibration, the credibility of the results obtained through SLIM can be enhanced by focused, detailed analyses for the calibration actions.

SLIM as applied in Beznau

The Beznau PSA is the earliest Swiss PSA study (1989/91 as surveyed) and the PLG SLIM implementation in this study lacks some of the refinements of the method as it is applied in the Mühleberg study. The main differences are:

- The SLIM implementation in the Beznau study uses fixed PSF weights in the summing of the different PSF influences as opposed to weights based on operator qualitative judgements.
- A single calibration scale is used in this implementation. In the later implementations, groups of similar actions are calibrated separately.
- The Beznau PSA does not use calibration data from other PSAs at all; instead, an undocumented correlation curve is used to convert the likelihood index on the relative scale to probabilities.

6.5.10 UK - Siz

The observations relate mainly to work done in support of Type C (dynamic) errors, and are based on an appraisal within the project of peer review comments, and developments being considered in connection with the Living PSA.

The approach to the assessment of dynamic errors was to model the required actions as three phases: diagnosis, decision-making and execution. The division into these three phases within an operator action (event) tree enabled within-task dependencies to be identified and assessed by the use of conditional probabilities, which is important if recovery routes or equipment unavailabilities are modelled within the task. This task division lent itself more to assessments by THERP than HEART, but the latter was applied in a conservative manner to the task elements, with the more detailed THERP approach only been employed where the significance of the task justified a more detailed approach.

The process of identifying Type C operator errors that needed representation in the PSA was driven more by an understanding of what operator actions were required to operate the plant in post fault situations, rather than by a systematic review of the procedures themselves. This was a consequence of the procedures being developed in parallel with the PSA, both of which were being served by Human Factors studies using task analysis. This illustrates one of the benefits of performing the PSA at an early stage of the plant development, which resulted in some Type C errors either being designed out, or a decrease in the need for certain operator actions (or reducing the likelihood of operator response failures, through enhanced instrumentation, indications, procedures and interlocks). While the Operator Error Assessments could be revisited at some future time to take account of the as-built design and the procedures validated for use in operation, the potential benefits from doing so can only be assessed after a period of commercial operation, and a review of risk contributions from the Living PSA. Any review could be selective.

The HEART technique applied during the project was based on the first version, as noted in the reference given in section 6 below, with an in-house extension using 'power factors' to take account of within

system dependencies. This version of the HEART manual gives limited guidance to analysts on a number of sensitive judgmental areas, for example, the degree of proportion of the maximum effect for a given error producing condition that is appropriate. A revised version of the manual can provide more guidance to analysts. However, the need for judgements in the Human Factors area will continue to be an important requirement, and the focus provided by HEART on ways of error reduction is considered to be of benefit in screening the majority of the required operator actions.

The treatment of operator error dependencies across functions was complicated by the use of functional fault trees, rather than use of conventional event tree / fault tree models. This meant that across-system dependencies had to be identified through searching multiple HEP cutsets, for which interpretation was then more difficult due to absence of the sequence of events leading to the cutset condition. This required a thorough understanding of the plant dependencies. It is intended that the Living PSA will adopt a conventional event tree/fault representation, which will enable scenarios involving multiple HEPs to be identified more readily by studying the structure of these models for where the HEP terms are located.

The nature and extent of the fault analysis undertaken for Sizewell B (and the Operator Error Assessments which form a part of it), arose from the need to support the design and licensing activities. However, extended time scales during which the work was undertaken (involving a number of analysts), and limitations in the tools available for the scope required, has probably led to some internal inconsistencies of application, in comparison with PSAs with more limited objectives and scope.

6.6 HRA-based Improvements of Design and Procedures

This section discusses improvements to plant design and operational procedures resulting from the PSAs. Many survey responses note that these changes may not be strictly due to the findings of the HRA. In any case, these changes all relate to operator-plant interactions.

More than 40 specific examples of hardware and procedure improvements are listed below. They are categorised based on their main direction. For instance, procedure revisions that result from the modification of hardware system are not considered in the count of procedure revisions. The automation of previously manual operator actions is accounted for as the installation of a new capability. It should be noted that these numbers are for illustration purposes only.

Modification type	No. of examples
• new procedures	8
• revision of procedures, technical specifications	18
• installation of “new” systems, automated capabilities	12
• modification of systems (including actuation logic)	12

Table 6-12. HRA-based Improvements Summarised (listed by PSA study code).

PSA Code	Description of Modification
Alm	Spain - Almaraz PWR The Almaraz PSA led to approximately 30 modifications in plant procedures as well as to modifications in the plant design. <ul style="list-style-type: none"> • Test procedures modified and addition of components to position verification checklists. This was particularly important for some potential common cause failures, especially for the Auxiliary Feedwater System. • The EOP for switching to the recirculation phase of Safety Injection was modified to include the manual action of closing the suction valves from the Refueling Water Storage Tank. Failure to close these valves leads to transfer of this water to the containment sump when the switchover is made. • The plant design was changed to automate the closure of the suction valves from RWST in this situation. • A valve in the injection path for High Pressure Safety Injection to the RCS cold legs did not have its power breaker inserted in normal operation. This has been changed and the valve hand switch in the control room is now placed in “pull out”.
AP600	Italy - General Electric AP600 (PWR) Not available
B1100	Japan - Standardized 1100 MWe class BWR <ul style="list-style-type: none"> • The ADS Logic has been modified.
Bez	Switzerland - Beznau PWR <ul style="list-style-type: none"> • Automatic initiation of the NANO backfitted system was added. The survey response describes this independent set of systems providing a third train of redundancy. • EOPs modified to include more guidance for operator actions to restore support systems of front line systems. • Procedures developed to cross-tie operational electrical buses to emergency buses as necessary. However, these cross-ties were subsequently forbidden on deterministic bases.
Bor	Netherlands - Borssele PWR <ul style="list-style-type: none"> • Minimum flow bypass lines and low flow alarms were added to the Low Pressure Injection pump to prevent pump failure due to deadheading in small-break LOCAs. Formerly, the operator needed to open several valves to the sump. • Implementation of staggered calibration policy for level sensors/transmitters. This failure mode was a dominant contributor to actuation failures of sump recirculation Several resolutions suggested in connection with operator actions for rapid cooldown (IS-LOCA and SGTR scenarios). These include <ul style="list-style-type: none"> • improved procedures for SGTR • installation of a redundant pressurizer spray function for the bunkered secondary side reserve supply system • addition of a trip signal for the high-head ECCS pumps at high level in the Steam Generator • installation of automatic 100 K/hr cooldown of the plant in SGTR In addition, system modifications to prevent IS-LOCA events are under study.

PSA Code	Description of Modification
Dod	Netherlands - Dodewaard BWR <ul style="list-style-type: none"> Automation of the operator action to inhibit Automatic Depressurization (to address especially scenarios where the time window is 2 minutes). The failure of this action was the driving contributor to ATWS sequences, which are very important contributors to CDF in this study.
Doel	Belgium - Doel PWR <ul style="list-style-type: none"> Early feedback of PSA results was used to improve plant and procedures. Two examples are given in the detailed treatments. Indication position alarms for valves were implemented to improve human reliability.
DRS	Germany - Deutsche Risikostudie-A and -B (German Risk Study) <ul style="list-style-type: none"> Addition of an automatic system to support the operator in recognizing the correct gradient for cooldown. Addition of capability to initiate emergency shutdown system from the control room. Development of an emergency procedure for primary and secondary feed and bleed (this decreased the CDF in high pressure sequences by a factor of 10).
HTR	Germany - HTR-500 (medium-sized gas-cooled pebble bed high temperature reactor) <ul style="list-style-type: none"> Periodic simulator requalification exercises for the diagnosis of abnormal events Symptom-based procedures Display annunciating need to open bypass valves of the decay heat removal (DHR) recirculators after Loss of Main Cooling System (LMCS) Plant exercises for post-diagnosis tasks; organization of performance of parallel tasks
LMFBR	Japan - Medium-sized loop-type liquid metal fast breeder reactor Not available
Loviisa	Finland - Loviisa PWR <ul style="list-style-type: none"> Development of new EOPs and improvements of EOPs Addition of control system and improved procedures to prevent the entry of non-borated coolant, improvement of dilution system Improved detection of VLOCA leakages outside containment and automatic isolation of some VLOCA leakages Backfitting for primary-to-secondary leakages Other areas of improvements: maintenance practices, control of valve positions, control of heavy load fastening to crane to prevent in-containment heavy load drops
Müh	Switzerland - Mühleberg BWR <ul style="list-style-type: none"> A depressurisation logic and hardware was added that is triggered by low Reactor Pressure Vessel level only. The original logic was activated by this signal in combination with the high drywell pressure signal. A scenario was identified in which the absence of the latter signal led to a delay in automatic depressurisation resulting in fuel overheating and damage. An "ATWS switch" was added to eliminate the need to repeatedly inhibit automatic depressurization (ADS) in ATWS sequences
P1100	Japan - Standardized 1100 MWe class PWR <ul style="list-style-type: none"> Post-accident procedures were developed for ISLOCA Post-accident procedures were also developed for SGTR to allow cooling through an intact SG even if the failed SG is not isolated

PSA Code	Description of Modification
P900	France - Standardized 900 MWe Framatome 4-loop PWR <ul style="list-style-type: none"> • Addition of an automatic system to avoid rapid dilution of the primary circuit • Automation of isolation of the letdown line in Loss of Heat Sink scenarios • Automation of make-up to the primary circuit for Loss of RHR scenarios • Improvements to SGTR and loss of feedwater procedures. (In this context, note that symptom-based procedures have been implemented since the PSA.)
P1300	France - Standardized 1300 MWe Framatome 3-loop PWR <ul style="list-style-type: none"> • Procedures and technical specifications modified to reduce potential for Loss of RHR in mid-loop operation. An automated water supply is added for accident mitigation • Implementation of automatics to prevent spurious dilution of the primary coolant.
Pick	Canada - Pickering A Risk Assessment CANDU <ul style="list-style-type: none"> • Addition to station emergency operating procedures of actions to take to connect reactor building to pressure relief duct in the event of activity release into containment unaccompanied by a significant rise in containment pressure. • Inclusion in daily panel check of status of various valves in emergency coolant injection system to reduce likelihood of valves being left in wrong state undetected.
PIUS	Italy - PIUS Not available
SBWR	Italy - General Electric SBWR Not available
Siz	UK - Sizewell PWR <ul style="list-style-type: none"> • Design changed to provide a diverse source of control supply to the Motor Switching Devices that supply motive power to the Motor-Driven AFW pumps. An additional operator action is required to change over the control supply and re-start the pump. • Technical specifications modified to ensure that the operator has at least one hour to recover in Loss of Residual Heat Removal (RHR) system scenarios. • Procedures modified to start the Battery Charging Diesel Generators earlier, in order to ensure or establish long-term availability of the electrical supply. • To address operator overriding of Engineered Safety Features (ESFs), a default position was adopted that the Reactor Protection System (RPS) automatic controls have priority over the manual control of ESFs. All exceptions to this default on grounds of operational flexibility were analyzed on a case-by-case basis.
TVO	Finland - TVO (Olkituoto) BWR Development or revision of procedures for: <ul style="list-style-type: none"> • connection of electrical power supply generator from a diesel generator of a neighbouring unit • manual depressurization of the reactor from the relay rooms • refilling of the ECC water storage tank and the condenser • residual heat removal to a diverse heat sink

PSA Code	Description of Modification
WS2/3/4	<p>Korea (R.O.K.) - Wolsong Units 2/3/4 Pressurized Heavy Water Reactor (CANDU 6)</p> <ul style="list-style-type: none"> • Procedural guidance for “isolation of one of two ECCS heat exchangers” and “EWS operation for ECC heat exchanger in the event of RCW failure” is being moved from “technical documents” to the main logic diagram of the Abnormal and Emergency Operating Procedures • Procedural guidance is being developed for “crash cooldown operation” in the Abnormal Operating Procedure for Moderator System/Moderator Cover Gas System/End-Shield Cooling System Failure. • Recommendations for the staffing of the Secondary Control Area (SCA, secondary control room originally intended primarily for seismic events) were developed. In particular, two operators will be dispatched to the SCA to operate the EWS or EPS during emergencies.

6.7 Conclusions from the Survey

A wide spectrum of HRA methods were represented in the responses. This includes: (a) Decomposition or Database Techniques (THERP/ASEP, HEART); (b) Time Dependent Methods (OAT/TRCs, HCR); (c) Expert Judgement Based Techniques (APJ, SLIM/FLIM).

In some cases the analyses were supported by simulator experiments. Only a few studies use a single method; in most cases several techniques were combined. The responses provide a number of insights concerning mixing and matching different methods, criteria in the choice of HRA overall approach, and views of reviewers and PSA users. Some of the observations on methods follow below:

- The choice of the most suitable approach may depend on the application. Thus, screening (conservative) approaches may be fully adequate when analysing a plant under construction or in the conceptual stage. On the other hand, for operating plants the implementation of more refined best estimate approaches is desirable and much more feasible. While the HEART method has not been widely used in nuclear PSAs, the experiences from a full scope application for a plant under construction are encouraging. An “in-house” extension of HEART was implemented in order to account for within-system dependencies. Focus provided by HEART on ways of error reduction is beneficial; the main difficulty lies in the sensitive judgements concerning the appropriate fraction of the maximum effect for a given error producing condition.
- Simulator-based models have definite merits but their applicability is still restricted, which makes it necessary to combine them with engineering/expert judgement. Areas where quantification with these methods can be problematic include the analyses of the failure to formulate the correct response, of rare situations, and of actions in time windows exceeding 30 minutes. Some methodological problems concern the aggregation of test samples (of the simulator data), the treatment of factors other than time, difficulties to simulate some situations (particularly during shutdown), and the transferability of simulator data to real accident situations⁴.

4. [Mosneron-Dupin, 1994] provides a detailed summary of French experiences from application of simulator-based approaches, with emphasis on problem areas.

- Techniques based on expert judgement are extensively applied in PSAs⁵. For the SLIM family of methods significant differences exist between SLIM/SLIM-MAUD and FLIM. Gradual refinements of FLIM are for example reflected in the implementations in the Swiss PSAs. Some key characteristics of SLIM/FLIM implementations contribute significantly to the validity and consistency of the results, independently of the theoretical basis of SLIM. These include: expertise, elicitation (group process, PSF weighting and rating processes), index formulation, and calibration (data sources, grouping of actions) [Dang and Hirschberg, 1995].

The survey performed for this task makes it clear that the HRA-related parts of PSA studies are particularly difficult to compare. Quantification approaches, that is, the selection and combination of HRA methods, and analysis assumptions, for instances, the crediting of recovery factors⁶, vary widely. In addition, however, both plant-specific and analysis-specific differences contribute significantly to the variability of the failure probabilities for the operator actions identified as important in each study.

The logic for the automatic initiation of systems is a key plant-specific characteristic with the potential to strongly influence the operator action models and HRA results. Four cases can be distinguished:

1. The action is manual; this action is not automatically performed in any circumstances (including other scenarios).
2. The action is a back-up to the (failed) automatic initiation of a system.
3. A successful response to the scenario requires that the operators anticipate the automatic initiation of a system. (In some scenarios, the automatic initiation will occur too late.)
4. The action is manual; however, the initiation of the system is automatic in other circumstances.
5. It can be seen that these four cases define essentially different operator actions that are not directly comparable.

Analysis-specific differences include the approach to modeling similar scenarios, in particular, whether operator actions are decomposed into separate actions (at the fault tree or event tree level) for the diagnosis/decision part and the execution part. This decomposition typically has no numerical effect on the core damage frequency (given appropriate and consistent consideration of dependencies); however, it does affect the importance measures reported for the operator actions (it reduces these) and consequently, their importance relative to other operator actions and to hardware.

A detailed look at the treatment of operator actions can be strongly recommended; one is able to identify the factors that drive the results numerically. In spite of similar numerical results, different methods may

-
5. Discussion of the applications of expert judgement in HRA and its use for analysis of organisational behaviour, as well as case studies are provided in [Reiman, 1994].
 6. Recovery factors refer to factors that may be credited to reduce a base human error probability; some examples of recovery factors are the availability of redundant indications or indication of the effectiveness of an action (feedback). They are distinct from recovery actions, which are separate operator actions to recover or substitute for failed equipment that are credited to reduce a cut set probability.

have been used or the same method applied differently; in addition, the definition of the action (as mentioned above), the credit taken for recoveries, and other assumptions may vary.

Finally, the HRAs have led to many useful plant and procedure improvements, examples of which appear in this chapter. Relatively few cases can be attributed directly and unambiguously to the HRA, which is an integral part of the PSA.

6.8 Chapter References

- /1/ Chien, S.H., A.A. Dykes, J.W. Stetkar and D.C. Bley (1988), "Quantification of Human Error Rates Using a SLIM-Based Approach," 1988 IEEE Fourth Conference on Human Factors and Power Plants, Monterey, CA, June 5-9, 1988.
- /2/ Chu, T.L. et al. (1994), "Evaluation of Potential Severe Accidents During Low Power and Shutdown Operations at Surry, Unit 1: Analysis of Core Damage Frequency from Internal Events During Mid-Loop Operations", Main Report, NUREG/CR-6144, U.S. Nuclear Regulatory Commission, July 1994.
- /3/ Dang, V. N. and S. Hirschberg, (1995), "Human Reliability Analysis in Probabilistic Safety Assessments: Current issues, the Swiss studies and options for research", prepared by Paul Scherrer Institute, Würenlingen and Villigen, for the Swiss Federal Nuclear Safety Inspectorate, HSK-AN-2887, December 1995.
- /4/ Dang, V.N. and S. Hirschberg, "Human Reliability Analysis in Probabilistic Safety Assessments: Current issues, the Swiss studies and options for research", prepared by Paul Scherrer Institute, Würenlingen and Villigen, for the Swiss Federal Nuclear Safety Inspectorate, HSK-AN-2887, December 1995.
- /5/ Embrey, D.E., Humphreys, P., Rosa, E.A., Kirwan, B., and Rea, K. (1984), "SLIM-MAUD: an approach to assessing human error probabilities using structured expert judgement," NUREG/CR-3518, U.S. NRC, March 1984.
- /6/ Hirschberg S., V.N. Dang, and L. Reiman (1996a), "International survey of PSA-identified critical operator actions", Invited paper, Probabilistic Safety Assessment and Management, ESREL '96 /PSAM III, Crete, Greece, 24-28. Jun. 1996, Vol. 1, 656-661.
- /7/ Hirschberg S., V.N. Dang, and L. Reiman (1996b), "A survey of international practice in HRA", Proceedings of PSA '96 International Topical Meeting on Probabilistic Safety Assessment, Park City, Utah, USA, 29. Sep.-3. Oct. 1996, Vol. 1, 78-84.
- /8/ Mosneron-Dupin, F., "Is Probabilistic Human Reliability Assessment possible?", presented at the EDF International Seminar on PSA and HRA, Paris, 21 - 23 November 1994.
- /9/ Reiman, L. *Expert judgement in analysis of human and organisational behavior at nuclear power plants*. Ph.D. Thesis, Finnish Centre for Radiation and Nuclear Safety, Helsinki, STUK-A118, December 1994.

7. SPECIAL TOPICS IN HRA

7.1 Modeling Errors of Commission

One of the major criticisms of current PSAs is that they do not adequately address an important class of human-system interactions, namely inappropriate actions, particularly those that might occur during the response to a transient or an accident, that place the plant in a situation of higher risk. This class of inappropriate actions is often referred to as errors of commission. This term has different meanings to different disciplines, therefore, in this section, we first define what we mean by errors of commission in the PSA context. We follow this with a discussion of some of the issues related to their inclusion in PSA models, and briefly describe an approach that has been used to address errors of commission in both a power and a low power and shutdown PSA. Work that is ongoing to develop more a comprehensive approach to the identification of errors of commission for inclusion in PSAs will be discussed in a later section.

7.1.1 Definition of an Error of Commission

The principal characteristic of an error of commission in a PSA context is that its consequence is a state of unavailability of a component, system or function. This is in contrast to an error of omission, which is characterised by a lack of action, and therefore preserves the status quo of a system, component, or function. Thus an important error of commission is recognised by its consequence. In the PSA context, the most significant errors of commission are those that either, in addition to resulting in failure to perform some function, also fail, or made unavailable, other equipment or functions needed to mitigate the accident scenario, or otherwise exacerbate the situation. Such an error introduces a dependency between events within a fault tree, or between functions on the event tree. However, an error of commission that fails a single function or system (e.g., by premature termination) can also be significant from a risk assessment perspective, if it provides a new mechanism of failure that cannot reasonably be expected to be included in the failure probability assigned for the function/system. This latter type of error does not introduce a new dependency between the events of the model, but increases the failure probability of the affected function.

As will be seen later, errors of commission may not necessarily result from an error on the part of the plant personnel directly involved in the activity. Instead, it is likely that they may have been set up to fail by the context in which they are asked to function. The terminology inappropriate actions does to some extent remove some of the connotation of blame that might be associated with the term error. However, because it has become common parlance, we will continue to use errors of commission.

It should be noted that some of the impacts of errors of commission may be incorporated in a PSA, even if they are not explicit in the structure of the logic model. For example, THERP [Swain and Guttman, 1983] includes contributions to the probability of failure to perform an action from such errors of commission as selecting the wrong switch from an array of switches. However, using the THERP approach, the impact of these errors of commission on other equipment is not modelled. The consequences of the errors of commission included in THERP models are essentially incorporated in the failure to perform a required function, i.e., an error of omission.

The modelling of errors of commission has historically been ruled as being out of scope in PSAs, probably because of a fear that including all possible errors of commission might lead to an increase in size of PRA models to such an extent as to make their solution impossible. However, recent progress in understanding the causes of human error has indicated the possibility of developing approaches to deal with this problem in a rational and economical way. To understand how this is the case, it is first necessary to recall how human failures are modelled in PSAs.

7.1.2 Modelling of Human-System Interactions in PSAs

Human Failure Events as Elements of PSA Models

The impact of human/systems interactions are included in PSA models by including events, which we will call human failure events, as basic elements of the logic models. There is a considerable variety in the way PRA analysts construct their plant logic models, although all have the common feature that inductive logic models called event trees are used to define the accident scenarios of interest. An event tree is a pictorial representation of the accident scenarios that can result from an initiating event. The scenarios are differentiated by the successes and failures of the events represented by the branch points in the event tree. These branches relate to the functions required to respond to the initiating event. The major discriminant between different approaches is in the detail which goes into the development of the event tree scenarios. Some analysts prefer to keep event trees fairly simple, with the branches of the trees representing successes and failures of critical safety functions; others prefer the branches to represent the successes and failures of the individual systems that can provide those functions; still others prefer that the branches represent successes and failures of trains of systems.

Other logic models, primarily fault trees, are used to model which combinations of component failures or unavailabilities lead to the functions, systems, or trains failing to perform their mission as specified in the success criteria associated with the event tree branches. The logic model events which represent the unavailability or failure states of the basic components are the basic events of the model. The basic events are the lowest level of decomposition in the model and therefore define the level of detail the model is capable of supporting.

Failures in human-system interactions are included in this set of basic events as events that represent specific failure modes of components, or of functions, that are a result of failures on the part of the plant personnel in their interactions with the plant. These HFEs may appear explicitly as event tree headings or they may appear in fault trees, depending on the impact of the failures, and the preference of the PRA analysts. The human-system interactions of interest may occur in the pre-accident, initiating event, or post-accident phases.

On The Definition of Human Failure Events

In a PSA, a detailed scenario description, obtained by combining the event tree and fault tree decompositions, is given as a product of an initiating event and a set of basic events, and each set forms an accident scenario cut set. The set, or a subset, of the basic events that appear in the same scenario cut set as a human failure event define, to the extent they can, the boundary conditions under which the human-system interaction should be evaluated. However, event tree models are constructed in such a way that several cut sets may be associated with one accident sequence. A particular human failure event may appear in many of these cut sets and is therefore associated with several different sets of basic events. Thus, for a single HFE, there may be several different conditioning scenarios. For example, one cut set associated with an HFE may contain among others, the event, 'pump A fails to start', whereas another cut set may contain the same events except that the fails to start. failure mode is replaced by 'pump A fails to

run'. These cut sets therefore represent a different plant condition, because, for example, the decay heat level is different. In fact the latter cut set is itself representative of a population of plant conditions, since a failure to run could occur at any time during the mission defined for the function for which the basic event is a contributing failure. Event trees can be, and are, used to capture some of the temporal and causal relationships between failures, but solutions to accident sequences are still generally not defined precisely in this respect. Usually, when it is necessary to fix some variable, for example, the time to onset of core damage, a particular realisation is chosen as a representative of the population of possibilities. In the above example, the time to onset of core damage would typically be based on the assumption that all cut sets lead to failure of the function at the time of the initiating event.

Not all relevant factors are captured in the models. As an example, instrumentation failures, which can adversely impact operator response, may not be explicitly modelled. In addition, it is also important to recognise that PSAs do not try to capture individual differences between, for example, nominally identical components, or between operating crews, but represent time averaged and population averaged reliability characteristics. These examples illustrate the ways in which the typical PSA scenario definitions do not explicitly address issues that might have an influence on identifying the error forcing conditions that might lead to errors of commission. Thus, an HFE, as it appears in a typical PSA, does not have clearly defined boundary conditions in terms of all the possible influences on operator behaviour.

There is a second important issue related to the definition of HFEs. Nuclear power plant systems are often forgiving in that the time constant of the system response often allows the operators to receive feedback and take corrective action before an irreversible change in plant status has occurred. In a PRA the transitions between plant states are modelled as being discrete, and therefore, there is an implied time period associated with the transition during which a recovery could be effected. Therefore, another important feature of human-related failures that must be addressed is the dynamic nature of the human-system interactions as described in the plan-execute-verify iteration discussed by Swain and Guttman [Swain and Guttman, 1983] among others. For example, an error may occur because of a slip on the part of a crew member, but it does not become a failure in the context of modelling the system unless it is allowed to remain uncorrected for long enough to cause an irreversible change in the status of a component or function. In other words, the HFEs that should be modelled in a PRA are generally not events occurring at a specific point in time, but represent failures of an iterative interactive process.

Conceptually, then, the human failure events in PSAs incorporate three phases of the human/system interaction, namely:

- the occurrence of the conditions that require a response on the part of the plant personnel,
- the initial committing of an error,
- failure to recover from the error before undesirable consequences occur.

A Quantification Model for an HFE

As discussed above, the interface between the modelling of human error and the system model is in the definition of human failure events as failure modes of functions or components of the system. The failure mode occurs in the model as a result of other conditions that require the function of interest to be performed, which is the first phase of the modelling of the error. Each error expression may be the result of one of several different error modes. An error mode is a phenomenological statement that enables an observer to recognise that an error has been made, by observing that the operator response is not what it

should have been. It is reasonable to suggest that each error mode may be the result of one of several different error causes, or error mechanisms, each of which may be activated by a different set of conditions. Furthermore, it is also reasonable to argue that the recovery from errors resulting from different causes are also influenced by different conditions.

This discussion suggests that one form of a mathematical model, or formalism, for the evaluation of the probability of an HFE, which we will refer to as an HEP, could be that given schematically in the following equation [Parry, 1995]. This formulation of the HEP model is expressed as an implicit model, in which the impact of recovery, in the sense implied by the previous discussion, i.e., at the level of the HFE, is incorporated into the event and its associated probability.

$$P(E|S) = \sum_{\substack{\text{mode } j}} \sum_{\substack{\text{mechanism} \\ i}} \sum_{\substack{\text{recovery} \\ \text{mechanism} \\ k}} P_{ji}(S) P_{nr}^{jik}(S)$$

where, $P(E | S)$ is the probability of error expression E in scenario S , $P_{ji}(S)$ is the probability of mechanism i resulting in mode j in scenario S , and $P_{nr}^{jik}(S)$ is the probability of non-recovery from mode j , mechanism i via recovery mechanism k in scenario S .

It is of course equally valid to treat the recovery explicitly as a separate event in the logic model. However, in this case the PSA logic model would have to be considerably expanded in size, and therefore, it is a much more attractive proposition to try to develop an implicit model.

7.1.3 Understanding Causes of Error

To make this a practical model, it is clearly necessary to focus the search for the most significant conditions and error mechanisms. This can be done by understanding how inappropriate actions can occur, i.e., what sets of conditions lead to enhancing the chances of errors.

Recent work in the behavioural sciences (see for example, [Reason, 1990] and [Hollnagel, 1993]), and the development of a multidisciplinary framework for the analysis of human/system interactions [Barriere et al., 1995], has contributed to the understanding of the interactive nature of human errors and plant behaviour that characterise the accidents that have occurred. This understanding suggests that it is essential to analyse not only the human-centred factors, with consideration of such performance-shaping factors as man-machine interface design, procedures content and format, and training, but also the conditions of the plant that both give rise to the need for actions and create a particularly challenging context. Examples of such conditions include misleading indications, equipment unavailabilities, and other unusual configurations or operational circumstances. This is in contrast to the existing HRA methods that consider principally the human-centred causes, with plant influences specified only as bounding definitions consistent with the associated PSA scenario definitions. The spectrum of possible plant conditions and the potential for encompassing particularly challenging scenarios has not generally been addressed.

Therefore, typical evaluations performed in HRA assessments of performance-shaping factors, such as the layout of indications or control switches, may not identify critical problems unless the whole range of possible plant conditions under which the controls or indicators may be required is considered. In other words, a particular layout of indicators and controls may be perfectly adequate for the nominal conditions

assumed for a PRA scenario. However, it is possible that there are other conditions that could be included within the same PRA scenario that would make the layout have an influence on the occurrence of operator errors in the accident response. For example, under during an accident scenario, an operator may be required to perform a series of actions at locations on several control boards. Provided the actions can be well separated in time, the layout may prove adequate. However, it is possible that under some subset of plant conditions for the same scenario, the dynamics of the plant require the actions be taken almost simultaneously. In this case the layout is inadequate and might result in failure to perform the actions in time.

Unless the analysis of PSFs is performed recognising that, as discussed above, plant conditions can vary significantly within the definition of a single PRA scenario, and that some of those plant conditions can be much more demanding of operators (both in terms of the plant conditions themselves and the limitations in PSFs like procedures and training under those conditions), the analysis may fail to identify the most likely conditions leading to operator failure.

Stated another way, operator failure associated with a PRA scenario is perhaps as likely, or more likely, to result from the "off normal" plant conditions that may represent one of the realisations of that scenario as it is to result from a random "human error" that might occur under the nominal conditions. Analyses of power-plant accidents and near-misses indicate that the influence of off normal plant conditions appears to dominate over "random" human errors.

Therefore PRA, to provide an effective tool for measuring and controlling risk, must be able to incorporate realistically those human errors that are caused by off-normal plant conditions as well as those that occur "randomly" during the assumed nominal accident conditions. However for PRA to incorporate errors caused by off-normal plant conditions, it is necessary to be able to identify under what conditions the human errors can be "forced", and to be able to estimate how likely these conditions are, and what are the likely consequences in terms of inappropriate human actions or inactions.

The identification of these error-forcing contexts must be based on an understanding of the kinds of psychological mechanisms causing human errors that can be "set up" by particular plant conditions that lie within the PRA definitions of accident scenarios. Without such an understanding, the search for these error-forcing contexts would be limited to searches for "repeat events" that were simply duplicates of earlier incidents where people had failed.

7.1.4 Approach to Analysis of EOCs

This section describes one approach that has been developed for the analysis of errors of commission and applied to both the full power mode and the low-power and shutdown modes of operation [Julius et al., 1995; Julius et al., 1996]. The method is intended to be applied to only rational actions.

The philosophy adopted in developing these procedures is essentially similar to that in [Macwan and Mosleh, 1994], namely a successive screening of the very large number of situations in which operators interact with the plant in order to find those cases where the likelihood of the scenario with the potential for leading to an unrecovered error of commission with potentially significant consequences is relatively high. The major difference is that, while the approach of [Macwan and Mosleh, 1994] relies on simulation to identify the opportunities for error, the approach presented by Julius et al. [Julius et al., 1995; Julius et al., 1996] is of necessity more economical and focused. The first step in the process is to identify those PSA defined scenarios that provide opportunities for a human/system interaction. The second step is, for each PSA scenario, to use models of error mechanisms and error causes to identify under what conditions, within the boundary conditions implied by that PSA scenario, an error of commission might occur. Once

an error is considered plausible, it may be screened out on the fact that the consequences of the likely errors are unimportant, or that it is very likely to be recovered, or that the likelihood of its occurring in the first place is sufficiently low.

In developing the method for full power modes of operation, it was assumed that the plant personnel are responding to an accident guided by procedures, and are trained in those procedures, which for the plant of interest were of the Westinghouse symptom oriented procedures. The procedures used by the operators trigger actions based on the status of equipment and by the trends in certain process parameters, such as pressure, level, power, and temperature. The procedures are used in the analysis to provide a means of identifying when operators are expected to perform certain functions and the cues they use to guide them. If the conditions are optimal, the likelihood of significant error should be minimal. Therefore, the philosophy adopted in developing the method was to search for scenarios in which conditions are not optimal. The errors were analysed in two groups; those which result in a wrong diagnosis of the accident type, and those which result in a failure of the execution of a response.

Misdiagnosis of Accident Type

A diagnosis of the accident type is made by the control room crew on the basis of the information available to them. It was assumed, therefore, in developing this procedure, that for an error of commission of misdiagnosis to be made, the crucial factor is that there be certain conditions associated with the scenario (e.g., instrumentation failure) that either distort the available information so that it looks like the signature of another scenario for which the inappropriate action would indeed be appropriate, or conditions are such that the operators are led to interpret the (correct) information incorrectly, e.g., by matching to a more familiar scenario signature. These assumptions are consistent with the similarity matching, and, to a lesser extent, frequency gambling processes that are proposed by Reason [Reason, 1990] to be the underlying processes that drive human cognitive behaviour. In this way, a model of error causes was incorporated. In addition, ambiguities, or unclear directions in the procedures (referred to as "type of response required") is considered at this stage. Making use of these ideas, not only can the potential opportunities for error be identified, but also the way in which the signature must be distorted to result in specific consequences can be determined.

Using the PSA model, the consequences of an error are relatively straightforward to identify once it has been identified. In addition, using the information on accident progression within the PSA, the opportunities for recovery before the consequences of an error becomes irreversible are relatively easily identified. The operator's expectations of the plant response to his actions, the strength of belief in his diagnosis, and memory of recent actions are all influential in the recovery process and were addressed as needed during screening at a later stage.

The identification of the opportunities for misdiagnosis of accident type was performed as follows:

1. Develop a Procedure Response Matrix (PRM) for all initiating events, or initiator groups that produce significantly different plant responses. This is a table of the expected trends of the important plant parameters and/or indicators identified in the E-0 procedure (e.g., Primary Pressure) for each of the initiator groups. Then, for each of these major groups, review the decision points in the procedural path. For each decision point that relates to entering a new path in the procedure (e.g., entering E-1, E-2, E-3A, E-3B, etc.), identify the possible incorrect decisions resulting from either misinterpreting or failure of the plant to provide the correct information used at the decision step, or missing the decision step altogether.

Screening can be performed on the basis of consequences, on the basis of procedurally guided recovery to the appropriate procedure, or on the basis of likelihood. The next step collects the information that will be useful at a later stage in the procedure.

2. Identify the critical indicators and alarms corresponding to the entries in the PRM developed in Step 1 above. Also, for each alarm and indicator identify:
 - Indicator/Alarm location(s)
 - Redundancy level
 - Whether these are diverse indicators or alarms (e.g., alternative methods of verifying the status of the critical parameter)

It is convenient to summarise this information in the form of a Plant Information Matrix (PIM), listing the critical parameters and the above information for each parameter, to be used in the screening at a later stage.

Screening on the Basis of insignificant Consequences

3. For each potential error, review the incorrectly applied procedure and identify whether there exists the potential for failing, not performing, or otherwise ranking unavailable, those functions required to bring the plant to a safe stable state (identified from the event tree). For those errors for which this cannot be excluded unequivocally, proceed to steps 4 through 6.

Screening on the Basis of the Potential for Recovery

4. For each global misdiagnosis opportunity review the procedure entered as a result of the misdiagnosis to identify re-diagnosis opportunities based on procedural directions. If there are such opportunities, the case can be screened from further analysis, if the conditions in 5 and 6 below are not met.
5. If the scenarios involve additional hardware failure over and above the initiating events, (e.g., failure of AFW), identify the impact on the operator's ability to make a re-diagnosis. This can be done by considering the impact of the hardware failure on plant response. Specify this in terms of changes in the values of the critical parameters in the PRM matrix that are used to make the re-diagnosis. If the recovery cannot be guaranteed, proceed to frequency screening.
6. Assess the potential for mindset leading to failure of the recovery action. This may, for example, be a function of training bias, or of a commonality of cues. For example, if the information used that resulted in the misdiagnosis in the first place is different from that that would be used to provide the opportunity for re-diagnosis, i.e., if common cues are used for the initial diagnosis and the re-diagnosis, do not claim recovery. If the recovery cannot be guaranteed, proceed to frequency screening.

Screening on Likelihood

To screen a scenario on the basis of frequency, it is helpful to consider different possible mechanisms that can be postulated that could lead to a symptom being perceived as different from what it should be:

- a) instrumentation failure, or unavailability of an indication due to equipment failure, or loss of support system,
- b) plant behaviour may be modified by other equipment states (e.g. , Steam Line Break inside containment may look like a small LOCA)
- c) the step in the procedure that completes the diagnosis may be missed, bypassed, or misinterpreted.

For (a) follow step 7 below:

For (b) follow step 8 below:

For (c) follow step 9 below:

7. Postulate the appropriate failure mode of instrumentation for the appropriate critical parameter for each case. Primarily focus on single failures of instrumentation; multiple instrumentation faults generally have a very low probability because of redundancy.
 - Check the possible failure modes (high, low, as-is) to determine if it is likely that the instrumentation would fail in such a way as to give incorrect information in the manner required to cause the error.
 - For each possible failure mode, using the PIM, determine the possibility of the operators recognising instrumentation failure, using information on redundancy of instrumentation, standard practices for checking alternate indications, procedural backup, etc.

Retain only those possibilities for which there is no obvious immediate backup to aid recovery from the incorrect information.

8. Identify potential equipment failures that can produce misleading indications by using the following:
 - Identify the status of equipment as it should be in response to the initiating event, e.g., PORVs remain closed, Safety Injection (SI) pumps should not start, etc.
 - Identify failures that result in violating these conditions and affect the key symptom(s) for decision-making in such a way as to lead to an incorrect decision.
 - Identify other symptoms or indications that are triggered by the additional faults and that can lead to an opportunity to correct this misinterpretation.

Retain those failures than can lead to misinterpretation, and allow no immediate potential for recovery, and which have a non-negligible probability of failure.

9. Assess the possibility of the operator not noticing, or misinterpreting the discriminating information available from the alarms and indicators. To do this list whether the following conditions exist for each initiator:
- Work overload (typically the number of alarms, and the number of parameters to be monitored for the initiator, on a relative basis)
 - Perception of time urgency based on training (this can be determined by interviewing operators and asking them to rank different initiators with respect to an overall urgency factor). The rate of change of the parameter is one way of measuring this.
 - Whether there is information supplied by an instrument that is known to be unreliable. This is a random hardware failure phenomenon, and the reliability of an instrument can be assessed as high or low for each of the important instruments. Persistent problems with instruments over a long period of time should be noted. This is a problem of desensitisation to information.
 - Whether there are any negative human factors considerations, e.g. , lack of clarity of information, remoteness of recovery equipment, lighting, etc.
 - Whether the procedural instruction is unambiguous and clear.
 - Whether training has over- or under-emphasised the scenario.

One potential way of using this information is to perform a qualitative screening as follows, using a rating scheme such as: High, Moderate, Low. A "High" rating on any of the first four factors or a high negative rating on the either of last two is translated into a "High Likelihood" of ignoring confusing information. Two or more "Moderate" ratings will also translate into "High Likelihood", and other combinations are screened out as "Low Likelihood".

For each alarm or indicator assessed as having a high likelihood of not being noticed by the operator, use the PIM to identify a recovery possibility, that is, the potential for the operator to eventually notice the information or learn about the plant condition through other means listed in PIM.

Errors in Response Implementation

It is the primarily the common cause potential of these errors which is of interest. The identification of the opportunities for errors to cause such failures was performed as follows:

1. For each initiating event group, list the functions appearing in the event tree and identify the set of success paths.
2. Identify possible human induced failure modes of the first function on a given success path, e.g., initiate system prematurely, terminate system prematurely, create diversion path, provide too much flow, provide too little flow.

3. Identify reasons for activating these failure modes by reviewing the procedure being followed to identify steps requiring some action as a result of a "result not obtained" statement, or a symptom- driven requirement to perform some action.
4. Each such step constitutes an opportunity for an error of commission if it satisfies the following:
 - a) changes the state of the system from that required
 - b) there is no automatic realignment
 - c) an override of an interlock is not necessary (the necessity to perform an override of an interlock is a powerful argument against the occurrence of a slip, though less so against a mistake)

The potential to recover from an error was assessed to be different if the error were a result of a mistake, i.e. , an error of intent, or of a slip, i.e. , an unintentional error. The assessment of the potential for recovery from a mistake in following the procedure, i. e. , from an incorrect plan was analysed in the following way:

1. Develop a list of potential error mechanisms based on the critical procedural step.
 - Incorrect information
 - error
 - indicator fault
 - Misread instruction
 - Omit a step in a procedure (this can cause an error of commission by omitting a precautionary action or a disabling action, for example)
2. Determine whether the error is irreversible - these are the most critical errors. For those errors that are reversible, using the information in Step 3, determine whether the operator may recover from misdiagnosis of the functional status of the system.
3. To assess the impact of potential immediate recovery or compensatory mechanisms, develop a System/Function Status Indicator list for each system which according to the above steps can be incorrectly operated. This is a list of alarms and indicators which provide the operator with the parameter values used in the procedures to assess the status of the system or function. In each case identify:
 - Redundancy level of indications
 - Diverse indicator or alarm (e.g. , alternative methods of verifying the status of the critical parameter)

- Procedural guidance or standard practice on checking and verifying the functional status of the indicator or alarm
 - Persistent alarm
 - Recoverable errors may be screened out.
4. If there are no immediate recovery possibilities, identify later recovery possibilities. This can be done by determining the plant or system response to the action and the availability of feedback information through indicators and alarms. Procedural steps should also be consulted for response guidelines. Errors, for which there is an opportunity for recovery, may be screened out.

The assessment of the potential for recovery from a slip error was performed as follows:

1. For the significant error opportunities, list important Operator Action Points (OAP) at the execution level (e.g., Open Valve x), identifying only actions essential to success.
2. Review control panel layout to identify the location of the switches associated with the action. Locate and identify other switches on the panel in the vicinity of those needed for action, and determine their function, and the likelihood of confusion, based on similarity of layout, etc.
3. Determine the criticality of the associated functions in terms of impact on the plant response to the accident. If there is no impact of accidental change in the status of these systems, the case can be screened out.
4. For cases not screened out on the basis of likelihood, postulate a Slip error for each of the functions in the vicinity of the intended function on the control panel (e.g. , open the wrong valve).
5. For each remaining case, list the possible system and plant response scenarios, and determine whether the error can be recovered using Steps 2 and 3 of the local misdiagnosis. The recoverable errors cases can be screened out.

Likelihood

The intent in developing this procedure was that, only if the error could not be screened out, was a likelihood to be formally assessed. The determination of the likelihood of an error is a function of the number and importance of performance influencing factors (PIFs) associated with the scenario generating the opportunity for error. In this procedure, the assessment of the likelihood was based on judgmental considerations, using PIFs that have been identified or postulated in previous research on causes of human error, and a formal quantification method was not developed.

Summary

In summary, the method consists essentially of the following steps:

- Identification of the human-system interactions that provide opportunities for errors to occur,

- Identification of the failure modes of functions, systems, or components that could occur from those errors,
- Identification of the most significant failure modes by:
 - Postulating error mechanisms, and
 - Screening the error mechanisms based on the existence and quality of defences against the error mechanisms, the likelihood of the error producing conditions, the consequence of the error, or the potential for detection and recovery.

Differences between full power and low-power / shutdown states

In full power models, it is primarily responses to upsets that are the source of errors. While pre-initiating human failure events are somewhat important but they are to a large extent independent of the post-initiating events, and are adequately analysed using THERP. They may cause complications in the definition of the EFC. However, while the procedure developed for the analysis for non-power conditions follows the same general approach as that for full power conditions, in detail it is different for several reasons. Firstly, in addition to addressing errors in the response to initiating events, it is more important in the non-power PRA to identify the errors that lead to, or contribute to the occurrence of, initiating events. In full power PRAs, errors resulting in initiating events are rarely modelled explicitly. This is because initiating events are usually plant trips, and the systems required to respond to the trip are, in most cases, different from those whose failures caused the trip. Failures in systems that both cause the trip and are required in mitigation of the trip are modelled as common cause initiating events. During non-power operations the initiating events are generally related to loss of a critical safety function (usually the loss of function of a system) rather than a reactor trip. In this case, there is often a dependence between the cause of the initiating event and the response, since the response is generally to recover that same function. Secondly, the operator responses are not as clearly guided by procedure as in the full-power case. For a Westinghouse plant, for example, there is generally no equivalent to the E-0 procedure to aid in the diagnosis of the event. Thus, misdiagnosis is potentially a much more significant concern. Furthermore, the diagnosis is more focused on identifying the root cause of the loss of a function and correcting for that root cause, rather than, as is the case of power operations, identifying a class of accident and responding by using systems intended to mitigate the effect of that accident. An additional consequence of the "root-cause-driven" response is a stronger dependence between what causes the initiator and how the operator responds to it. Thirdly, the plant configuration is constantly changing during an outage, and there are many different activities proceeding in parallel, with many more opportunities for plant personnel to interact with the plant. In the full-power case [Julius et al., 1995], the activities of interest were primarily associated with the control room crew responses to initiating events.

7.2 Modeling Dependencies

There are several ways in which the issue of dependency can arise in the analysis of human/system interactions. Firstly, it is widely recognised that the values of the human error probabilities (HEPs) are conditioned by, and therefore dependent on, the scenario with which they are associated. This type of dependence has always been the focus of HRA; it is in the choice of scenario characteristics that affect the probabilities that the various methods differ. There is a second type of dependence related to the modeling of the process that leads to the evaluation of the HEP. In a THERP analysis of an HEP for example, it may be considered that the operator's execution of two or more of the tasks required to successfully accomplish the response whose failure is represented by the HFE for which the HEP is being evaluated, might be such that, if he fails to perform one of the actions, he is more likely to fail the others also. This type of

dependence is adequately treated by techniques such as THERP. The third type of dependence, and the one that is of concern here, is that between the HFEs that occur together in an accident scenario definition. These HFEs may at first sight have no obvious dependence since, for example, they may refer to different functions. This type of dependence has been treated, if at all, in an ad hoc fashion, and yet it is extremely important, since ignoring this dependence can lead to a major underestimation of the scenario frequency.

7.2.1 Treatment of Dependence Requires an Understanding of Causes of Human Failure

A proper treatment of dependence requires an understanding of why the human caused failures can occur. Understanding the causes of human failures has not been a focus of HRA methods to date, which perhaps explains why the treatment of dependencies has been, as indicated above, somewhat ad hoc. The following example illustrates why a lack of understanding of causes of failure leads to problems with quantification.

Consider the following scenario in a specific BWR (adapted from Lydell and Parry, HRA and the Modeling of Human Interactions, Proceedings of PSAM, Beverly Hills, CA, 1992). The initiating event is chosen to be an MSIV closure. Since the plant has steam driven feedwater pumps, this leads to a loss of feedwater, and inventory make-up has to be provided by one of the ECCS systems, HPCI, LPCI, LPCS, or by RCIC, condensate, and/or CRD. The RCIC, HPCI, and CRD systems can inject at high pressure, the condensate system at an intermediate pressure, and the other systems at low pressure. If high pressure injection systems fail, it is necessary to depressurize the RPV. On a functional event tree, the first event following questions regarding successful scram, and operation of the safety-relief valves, is high pressure injection, which is assumed to have failed. The next event represents the depressurization function which is represented by an event X. While the automatic depressurization system (ADS) will automatically depressurize the reactor, the procedures are written to try to avoid exceeding the maximum cooldown rate by providing instructions to inhibit the ADS function, and allowing the operators to manually control the depressurization.

If the operators fail to inhibit the ADS, eventually, as the level drops, a blowdown will occur, and the low pressure systems will inject. Thus failure to inhibit is a safe mistake. If ADS has been inhibited, the procedures instruct the operators to proceed to emergency blowdown if the level cannot be maintained above the top of active fuel (TAF). What is significant, is that the action to inhibit ADS places the plant in a less safe condition for this scenario as it requires additional action by operators. Failure to initiate the blowdown by the time there is substantial core uncover will lead to core damage.

In discussion with the operating staff however, it was discovered that training is such that, if the condensate pumps are available, the operators will depressurize manually to bring the pressure to below the shutoff head of the condensate pumps. There should be no reluctance to do this as there is no penalty: the maximum cooldown rate will not be exceeded, and the HPCI and RCIC pumps could still be used if recovered in time.

Based on these considerations a cutset equation for the function X was derived as:

$$X = [CP + OA1]*OA2*[ADS] + [CP + OA1]*OA2* OA3 + HADS$$

where CP represents the cutsets for failure of the condensate pumps, ADS the failure of the ADS function, and HADS the failure of the hardware required to open an appropriate number of SRVs.

The OAi events are the HFES. OA1 is the failure to initiate early depressurization, OA2 the failure to inhibit ADS, OA2 is the successful inhibit of ADS, and OA3 the failure to initiate late depressurization.

The quantification of the cutset [CP + OA1]*OA2*OA3 can be used to illustrate an important point. Consider first the cutset CP*OA2*OA3. If the condensate pumps have failed, the operators will probably not try to depressurize early. Thus OA3 can be regarded as a simple response with the cue being the level reaching TAF.

Consider the cutset OA1*OA2*OA3. In this case, the operators have, for some reason other than failure of the condensate pumps, decided not to initiate early depressurization. The quantification of this cutset is problematic because we have no clear idea of the reason why they should have decided not to depressurize. The occurrence of the successful ADS inhibit does strongly suggest they are in the right procedure, and thus, it could be assumed that this cutset has a low probability.

7.2.2 *Methods for Accounting for Dependence*

Dependence Induced by Time Constraints

When multiple HEPs appear in the same scenario, one type of dependency that can be addressed by certain HRA models is that caused by the restriction caused by the time window. For example, if the total time window for a scenario is T_w , and the first action takes T_1 , then the time available for a second action is $T_w - T_1$. If a model for HEP is given as a function of time, then the probability of the product of the HEPs can be evaluated as a convolution integral:

$$P\{HEP_1 \cdot HEP_2\} = \int_0^{T_w} P\{HEP_1(T_w - T_1)\} \cdot P\{HEP_2(T_w - T_1)\} dT_1$$

Thus a TRC model can be used to capture this type of dependence. However, the HFES are still treated as conditionally independent.

Dependence Caused by Common Influences

It is generally considered that the most important aspect of dependency is that there must be some common factors that result in leading to failures of the multiple actions. These common factors can include common influencing factors such as procedural instructions, common indications etc. For example, it may be that the actions involved are all associated with the achievement of the same function, e.g., decay heat removal. There will be a dependence created by the reliance on a procedural directive that addresses the need to perform decay heat removal.

Consider an example taken from a PRA for a Westinghouse PWR with standard Westinghouse Emergency Operating Procedures. The initiating event is a loss of feedwater, necessitating a need for establishing some sort of heat removal. The events associated with establishing some sort of decay heat removal occur in the event tree in the order: AFW (auxiliary feedwater), MFW (main feedwater), HPI (high pressure injection), OFB (initiation of feed and bleed), for the case with no safety injection, and in the order: AFW, MFW, OFB for the case where safety injection has occurred. Each of the events MFW, HPI, and OFB are modeled by fault trees containing human failure events that represent failure to initiate the corresponding functions. These HFES are considered to be dependent by virtue of their being called out in the same procedure, FR.H-1.

One solution to modeling this situation is to model the cognitive contribution to the HFE, i. e. , the recognition of the need, and decision to perform the appropriate response, separately from that of the execution part. Further, in some PRAs, for this particular case, the cognitive part was separated into two contributions. A cognitive error was calculated for the failure to enter FR.H-1. This event is common to the functions MFW, HPI, and OFB. Another cognitive error was calculated for the failure to realise the need to initiate feed and bleed. Further, for the latter, two cases were evaluated; one for the scenarios when SI has not initiated, and another for the scenarios when SI has initiated. This type of modeling associates the dependency with specific human functions, and when the dependent part is identified it is complete.

In other examples of the treatment of dependency, a factor is applied to the cut set to account for the dependence. While some of these are considered judgementally, it is preferable to try to generate some guidelines that can be used to try to provide an auditable basis for the allocation of degree of dependence. Examples of such guidelines that have been used in previous PRAs are:

If operator actions are separated in time, and have different cues, and in particular if they are to be performed by different crew members, the HEPs can be regarded as independent.

If two HEPs are associated with actions which are taken in the same time frame and directed by a common procedure reader, they should be considered as potentially highly dependent.

A memorized action by one of the board operators is independent of another action performed by other crew members.

HEPs which represent failures in responses for which the same indications are crucial cues should be regarded as dependent, and the degree of dependence should be determined by the degree to which other, diverse indications are used to supplement the common indication.

When, in a sequence, the failures in responses are separated chronologically by successes, they may be regarded as independent.

Swain and Guttman, in THERP, give rules for assessing conditional HEPs given varying degrees of dependency, and these have been used fairly extensively, even though this was not the original intent of the rules. Use of the rules gives a degree of traceability to the analysis.

The development of improved HRA methods that directly address the causes of error should improve dramatically the treatment of dependency.

7.3 Modeling Recoveries

7.3.1 *The recovery issue*

Within the large flora of existing national and international guidelines on conducting and reviewing PSAs treatment of recoveries has not been given the attention it deserves. While there is no detailed standard for carrying out HRA, a set of generally accepted rules exists and is followed by most practitioners. Recoveries, or at least some types of recoveries, constitute an exception in this context. The degree of taking credit for recoveries varies extensively between the PSAs and in some cases amazing numerical

differences between the estimated core damage frequencies for similar plants were found to stem from this issue. The situation is complicated by the fact that there are significant analyst-to-analyst difference in what is regarded as a recovery.

7.3.2 *Types of recoveries and their treatment*

Recoveries involving human actions may act on different types of human errors and may appear in a PSA on different levels of the analysis. In addition, there exist different practices with respect to at what development stage of a PSA the recoveries are being implemented in the overall model.

Standard treatments of recoveries of latent maintenance failures, such as a safety-related valve left in the wrong position, exist and are relatively well supported by data based on operational experience [Swain and Guttman, 1983]. Distinction is made here between direct observation of deviant conditions through operations personnel via displays or other visible indications of equipment conditions, annunciated indications, and detection opportunities outside of the control room. Appropriate credit may be taken for human redundancy and availability of written procedures. Annunciators are considered a powerful recovery factor but too many annunciators may overwhelm operators.

Explicit modelling of the above recoveries is normally explicitly carried out whenever task decomposing methods (such as THERP) are used. For other types of methods, for example those based on expert judgement (such as SLIM), the credit taken for recoveries of this type is implicit and, consequently, much less transparent. The survey of HRAs, covered in the preceding chapter, illustrated this difficulty; it has not been possible on the basis of the available information to establish to what extent the methodological differences drive the differences in the treatment of recoveries, since the separation of the latter from the overall modelling of specific actions is not always possible.

Restoration of failed systems, based on procedural guidance and supported by training, are normally credited in PSAs. A typical example is restoration of offsite power. Other cases, where credit frequently but not generally is taken for restoration, include loss of DC buses and diesel generators. The most controversial recoveries are restorations and repairs of failed systems, and improvised actions, not supported by procedures. Different levels of crediting for such actions may lead to dramatic differences in the estimated core damage frequencies. These recoveries are normally implemented on a cutset-specific basis since their feasibility must be judged in light of the detailed conditions prevailing in the given scenario. It is worth noting that since recovery factors are frequently only applied to dominant cutsets the result may be a distortion in the relative ranking of contributors.

Good basis for generally taking credit for repairs under accident conditions is frequently missing. The repair time given in typical data sources usually only reflects the effective repair time. In reality, the repair task consists of a number of logical sub-tasks [Hirschberg et al., 1989]:

- disclosure of system breakdown
- tracing of failure from the central control room
- assembling the repair force
- preparation for repair (spare parts, tools, radiation protection, transportation)

- localisation of component failures and actual repair (this sub-task corresponds to the effective repair time)
- system restart

Experience shows that depending on the circumstances the effective repair time is not necessarily a totally dominant part of the whole repair. Furthermore, repair of failed systems may represent one strategy of dealing with the accident, competing with other approaches. This may lead to conflicting goals.

In conclusion, there is a need of harmonisation of recovery modelling within PSAs. This regards especially the level of credit that it is reasonable to take for repairs or improvised actions. Also the transparency of recovery modelling needs to be improved.

7.4 HRA in the Context of External Events Analysis

7.4.1 Background

The importance of external events⁷ is reflected in the results of numerous PSAs. Particularly in the case of older plants having lower degree of redundancy and separation, and normally less extensive built-in protection against external and internal hazards, such events frequently belong to the dominant contributors to the core damage frequency.

The type of human actions that need to be undertaken as a response to an external event may be event-specific. Thus, in the case of an internal fire the plant staff may need to: (a) undertake actions to mitigate the fire itself, and (b) to respond to the internal initiating event caused by the fire. On the other hand, seismic events as such can not be mitigated and only the second type of response (b) applies in this case.

The operator response to external events may be subject to specific difficulties, related to the characteristic features of such events. First, external events constitute Common Cause Initiators (CCIs), i.e. the redundant equipment needed for the mitigation of the event might have been disabled by the occurrence of this event. This complicates the situation and may lead to requirements on relatively short response times. Second, the information normally available to the operators may be distorted due to the impact of external events on instrumentation and signal processing. Third, the staff can be physically affected by the external event (e.g. by smoke). Consequently, appropriate modelling of human behaviour under conditions associated with external events is a complex task. Scarceness of relevant data, in most cases practically non-existent operational experience of situations characteristic for conditions that may appear upon occurrence of an external event, and limitations in simulator training to represent such situations, are additional factors contributing to the large uncertainties in HRA.

7.4.2 Current treatment

Internal fires and seismic events are commented on in the following. Among external events they typically constitute the most dominant contributors and particularly in the case of fires modelling of human actions is quite demanding.

7. External events include both natural hazards external to the plant (e.g. external floods and fires, tornados, earthquakes) and internal hazards (e.g. internal floods and fires).

Human actions in response to internal fire involve [Gil, 1996]:

- Detection
- Diagnosis
- Selecting the extinguishing strategy
- Hand extinguishing
- Performing in the time available

[Gil, 1996] provides a systematic structure for addressing these elements. For example, human detection can be direct (through actual presence or through inspection), or indirect (through generation of an internal initiating event or through damage to equipment). With respect to direct detection, a decisive factor for probability estimation is the attendance in the areas of interest; normally no credit is taken for direct detection in areas that are seldom attended. In spite of relatively large experience with industrial fires, generalisations of extinguishing times, an important parameter for the development of fire accident scenarios, are difficult due to the strong dependence on the type of extinguishing equipment and agent, fire size and character, accessibility to the fire area etc.

In the context of the response to an initiating event caused by a fire, transient-related accident sequences identified in the internal events analysis may require modifications, including need of further operator actions. Such changes can for example be necessary when the fire affects equipment whose failure in the course of accident propagation has not been taken into account in the internal events analysis, based on exclusion due to low probability. In any case, even if the structure of some accident sequences remains unchanged, the performance shaping factors may be totally different in connection with fire. The specific influences of the fire may concern:

- Loss of equipment as the direct effect of the fire (flames, heat, smoke), as a result of the use of extinguishing equipment, or due to chosen strategy to deal with the fire (e.g. tripping equipment, realigning trains). In particular, cognitive tasks may be hindered by fire-induced failures of alarms, instrumentation channels, status lights, controls, lighting systems, communication systems etc.
- Exposure of the staff to smoke and high temperatures. Extreme difficulties may arise if the fire occurs in the control room (e.g. cabinet fire). Other scenarios may include local actions that require access to areas directly affected by the fire.

Consideration of these influences is necessary when judging the feasibility of specific operator actions and estimating the corresponding probabilities.

Valuable experience, supporting the analysis and assisting in resolving some of the difficulties mentioned above, is available in the EPRI Fire Events Database [Parkinson et al., 1993; Oehlberg et al., 1994] and in related studies. For example:

- All experienced control room fires indicate that suppression occurs quickly as a function of time.

- The probability of smoke obscuring the main control board appears to be very low.
- The amount of smoke accumulating in the control room and originating from other source areas appears to be small.
- Fires in electrical cabinets are unlikely to damage other equipment.

Most current PSAs provide a rather rough treatment of human performance in connection to seismic accident scenarios. Operator actions required to mitigate post-seismic plant events, for the most part, are the same as those identified for internal events. However, similar to internal fires seismic events may lead to extremely high stress levels, confusion and difficulties to access areas outside of the control room. As a result the probability of the operators' success may change.

Current practice (e.g. [Chung et al. 96]) involves adjustment of probabilities estimated in the analysis of internal events by consideration of the available time (short, medium or long period) and location for executing the required action (in control room or ex-control room). In [Chung et al. 96] the factors used for these adjustments range between 1 and 30, depending on whether the combination of the performance shaping factors is favourable or unfavourable.

In most studies no or little credit is taken for operator actions during the first 30 minutes after a large earthquake [Budnitz, 1997]. At high seismic intensity levels, high stress conditions may, however, remain for a long time. Given the possibility of equipment damage, blocked pathways, jammed doors, injuries of fellow plant workers etc. the uncertainties associated particularly with ex-control room actions are well beyond those normally characteristic for operator actions associated with internal events.

7.5 HRA in PSAs for Low Power and Shutdown Conditions

In the French PSAs as published in 1990, the important insights were that :

- shutdown situations have a significant contribution to the overall core melt frequency ;
- during shutdown the role of human factor is dominant since generally there are no automatic device for actuation of the safety systems, the time windows available for operator intervention can be very short, the information available (alarms, measurements and the emergency procedures are often very limited).

In the French approach for probabilistic human reliability assessment (PHRA), a large emphasis was devoted to the use of simulator experiments. However, this very important source of information had important limitations in case of shutdown situations : there were no test data which could be used directly, and 'extrapolation' of models established with full-power tests need a difficult part of judgement since several specific characteristics of shutdown situations were not simulated. For example :

- several human actions can contribute to the same sequence, dependencies have to be assessed;
- activities outside of the control room (local actions) play an important role;
- during mid-loop operation, the operators have to realise a water-make-up in the primary circuit in several accident sequences (excessive draining, LOCA...). In this case there is a possibility of workers being present in the reactor building and especially inside the steam

generators (for fitting the nozzle dams for example) : this situation is a difficulty for decision making, there is a conflict between safety and personnel security which is not possible to simulate.

For all these reasons the PHRA quantification was generally performed with conservative assumptions, and the uncertainties are larger than for power states.

Presently an updating of the PSA's is in progress, especially for shutdown situations, since several changes have to be taken into account :

- automatism have been implemented on the plant, especially an automatic water make-up in case of loss of RHRS ;
- emergency procedures related to shutdown states have been improved ;
- several other improvement for helping the operators are introduced (alarms, level measurements, technical specifications, training...).

These modifications will reduce the contribution of shutdown state and of the corresponding human actions. However there are still difficulties for simulating these situations, and PHRA quantification during shutdown remains a problem. A way of research is presently a study of real operating experience, in order to validate the models.

7.6 Consideration of Organisation and Management Factors

7.6.1 Methodological approaches

In this section, two approaches for incorporating organisational factors in PSAs are summarised; these are the Work Process Analysis Method [Davioudian et al., 1994] and an influence factor approach [Moiene and Orvis, 1994]. To our knowledge, these methods are proposals and have not been applied generally in a PSA study. Unlike these methods that extend the PSA, Paté-Cornell and Fischbeck illustrate the use of the PSA as a tool for identifying key organisational factors (as well as human decisions and actions) that influence the risk. Their case study and approach does not attempt to quantify the influence of organisational factors.[Paté-Cornell and Fischbeck, 1993]

7.6.1.1 Work Process Analysis Method (WPAM)

The Work Process Analysis Method (WPAM) approach focuses “primarily on capturing the common-cause effect of organisational factors on parameters such as equipment failure rates” (p.080/7 [Davioudian et al., 1994]). To do so, it analyses work processes, the sequences of tasks designed within the operational environment of an organization to achieve a specific goal. In the first part, WPAM consists of a mostly qualitative analysis of work processes and assesses the importance of organisational factors in the overall quality and efficiency of the work process. In its second part, WPAM consists of a quantitative analysis of reach dominant accident sequence, whereby the effect of organisational factors is measured and incorporated into PSA results. It is worthwhile to note that the quantitative bridge between the organisational factor ratings and PSA parameters is based on an expert judgement process similar to that of SLIM. A paired-comparisons approach is used to generate the weights of the organisational factors in a likelihood index [Davioudian et al., 1994].

7.6.1.2 *APG's influence factor approach*

In [Moieni and Orvis, 1994], Moieni and Orvis propose an approach to incorporating organisational factors by extending the PSF (performance shaping factor) concept “to include organisational influences in HRA models as higher level influences that may concurrently affect several of the ‘traditional PSFs’ as well as introduce other direct influences on personnel performance” (p.080/1) In contrast to WPAM’s consideration of other parts of the plant organization, e.g. maintenance, this approach focuses on control room crew reliability. The influence of organisational factors and the quantification of their impact is based on the “use of decision trees, expert judgement, empirical data on human error (if available) and information collected on organisational factors in the form of ratings.”

7.6.1.3. *A PSA-centered approach*

In [Paté-Cornell and Fischbeck, 1993], the authors remark that there are at least two approaches “to link the managing organization to the performance of a system.” “A ‘top down’ approach starts with the organization and derives (or observes) from its characteristics its probable effects on system reliability. A ‘bottom up’ approach starts with the technical system and attempts to identify systematically the behavioural and organisational factors that affect the performance of each element.” (p. 241). Using this distinction, both WPAM and the influence factor approach would appear to be top-down approaches. On the other hand, the approach to considering organisational factors and risk discussed in [Paté-Cornell and Fischbeck, 1993] is bottom-up.

We choose to refer to this approach as PSA-centered because it starts with the technical characteristics of the system and the PSA inputs and results in order to identify the most risk-significant human decisions and actions and organisational factors. Furthermore, it shows that for some applications of the PSA, organisational factors can be systematically considered and prioritised on the basis of risk without attempting quantification. In a note, Paté-Cornell and Fischbeck also discuss an approach to computing the benefits of improving some of the organisational factors “that are the roots of the TPS [thermal protection system of the space shuttle] failure risk.” It seems worthwhile to note two apparent characteristics of this quantification. First, it appears impressively complex. Second, and perhaps more importantly, the relationships or links between the organisational factors and the risk analysis model are application-specific.

7.6.2 *Data issues*

Perhaps the single most important question concerning the study of organisational factors in the context of the nuclear power industry is, “What is the relationship between measures of organisational factors and plant performance?” Since plant performance is affected by equipment availability and reliability and human reliability, then the question becomes, “How can we investigate organisational factors issues such that we can define the correlation between measures of organisational factors and plant material condition and operation?”

The literature contains various theories of organisational structure and functioning, as well as numerous instruments for measuring organisations on “factors” which characterise organisations. It is essential that these factors be shown to be correlated with plant safety performance; without knowledge of such correlations, organisational factors measures will not have any meaningful predictive or direct use in managing (or, for that matter, regulating) nuclear power plant safety.

The question that research can address include:

- What are the appropriate measures of plant performance? Are they data such as equipment reliability and availability and human reliability that can be incorporated into probabilistic risk assessments, are they performance indicators such as the number of safety system activations, or a combination of the two?
- What are the plant practices that impact these measures of plant performance? In this context, plant practices include maintenance, operations, and design (including configuration management) activities. What does the analysis of operating data tell us about the effect of “good” and “bad” plant practices on plant performance? What is the nature of the impact of plant practices? That is, what is the impact of, for example, an above-average inspection program on the reliability and availability of a specific category of plant component, such as motor-operated valves?
- What are the organisational factors that impact plant practices? What does a given rating on an organisational factor such as “safety culture” mean in terms of plant practices? Are there “clusters” of organisational factors that interact in ways important to plant performance? What is the nature and magnitude of the effect the factor has on plant practice and, ultimately, on plant safety?

It seems that the best way to understand the influence of organisational factors on equipment and human performance is to first understand the effect of such factors on how the plant is run. That is, it seems that the first step in understanding the impact of organisational factors on plant safety is to understand the impact (that is, both the nature and magnitude of influence) of organisational factors on how plants are operated and maintained. This relationship is not presently well defined, and research in this area would greatly advance the current state of knowledge. Better understood, but still worthy of further inquiry, is the relationship between organisational factors and plant practices.

Perhaps an anecdote will help make the point. A gentleman associated with the computer industry related a story that told of how his company was trying to define the necessary ingredients of a good computer programmer. He said that if the company had made up a list *a priori* of the characteristics of a good programmer, then they would have had a hard time deciding which ones were more important than others. So, if the company had then used that list to screen applicants for computer programmer positions, they could easily have selected applicants that, although they had most of the ingredients, did not have the necessary ingredients. Rather, the company studied the good programmers and identified those characteristics that were common to all good programmers. Thus, they felt they had identified a smaller set of characteristics that were more likely to be well correlated with good performance.

The point of the anecdote above is that there is no substitute for research that investigates the basic relationships between variables, in our case, organisational factors and plant performance. Only with the knowledge of these relationships in hand can we reach conclusions regarding the importance of given organisational factors on the safety of nuclear power plants. The ability to conduct research to define these relationships will require access to data from a variety of sources, including individual plants, aggregate data bases, and expert panels.

7.7 Transferability of Simulator-Based Data

7.7.1 Introduction

To take into account and study human factor during the operation of a nuclear unit under accident conditions, full-scale simulators must be used. Realistic scenarios can be simulated and worked upon by the operator crews. Thus, data on the operator's behaviour can be derived from these tests and used as a basis for assessing new operator aids, for validating procedures for incidents and accidents conditions as well as for probabilistic safety assessments.

However, if simulators are extensively used to study exceptional situations, one should wonder whether the operator's behaviour observed during tests is representative of their actual behaviour. In other words, would the operators act the same under real operating conditions or, if not, is it possible to identify and quantify discrepancies between what happens in reality and on simulators?

7.7.2 Simulation biases

Whatever the type of tests performed, the operator crew is obviously never placed in perfectly realistic conditions, even when the simulator is said to faithfully recreate the physical phenomena.

These deviations from the real unit conditions are called biases. Apparently, the following major biases have been identified :

1. The simulated incident does not fit in an operating « history », in a context that may influence the operators.
2. On simulators, the operators are expecting the « worst » because they come to manage an accident, whereas, in real life, they expect what is most likely and, hence, what is most common. Then, they have a tendency to minimise the gravity of real events.
3. The operators may experience a certain stress, especially when they participate in their first test (they do not know « what will happen », they are not used to being watched, particularly when there are several observers who take notes without speaking) or when things turn out badly. But they are not under an as heavy stress as during a serious accident occurring in real life.
4. The operators are not much disturbed by « the outside world » (various phone calls, problems not directly pertaining to the accident recovery actions).
5. As the simulator recreates the control room only, it is very difficult to take actions to be performed outside into account.

For these last points, there is no representativeness of the management of the auxiliary operators by the crew, concerning for example the transmission of information between the supervisor and the operators (retransmission of the local actions report), disturbance, interruption in the control room's activity by the auxiliary operators, availability of them.

In return, they don't get additional help which they could find in real situations (additional operators, technical support...).

6. The operators do not work in their own surroundings. There are always some differences between the simulator and their actual control room (in their units, they often add some reading aids on the indicators ; they are accustomed to the ink colour used for a given parameter on a multitrack recorder... Though they look trivial, these differences may result in serious errors).
7. Each test is a training opportunity. There is, therefore, an evolution in the crew's behaviour from one test to another.
8. The operators are influenced by the observer, even if the latter do not intervene. For instance, the operators are probably induced by the observers' presence to conform to operating rules more than they would do in their plants.
9. The operators on the simulator do not act under economic constraints (power generation, equipment preservation). The safety requirements, therefore, overshadow economic constraints more than in an unit.
10. The operators feel compelled to act by the simulation context. Indeed, it is difficult to remain doing nothing when one has come especially to act and when one is observed, even if the need for action is very present on real situations too. For instance, unable to bear the long waiting (20 minutes) before the criteria for action were reached, several operators have made a reading error inducing them to think that the criteria were satisfied. On the other hand, in real situations they are more likely focused on material recovery non-existent on full scale simulators.
11. Because they know that on the simulator their errors have no serious consequences, operators may be tempted to act more rashly and less thoughtfully.
12. As a matter of fact, whereas general physical phenomena are well simulated, the component deterioration resulting from errors is often inadequately simulated.
13. The tests are generally carried out in the daytime, although it is well known that night has a significant impact on performance.

7.7.3 *Limitation of simulation biases*

During real-life tests, preventive measures are taken to limit biases and, in particular, those mentioned in 5, 6, 7 and 8.

For the points 3 and 8, there are previous talks with the operators to win their confidence to introduce the observers to them and to explain the framework of the program. That represents a significant difference with the tests on full scale simulators for training and validation of the procedures.

For the point 7, the tests don't last for more than three days for a team.

However, it is impossible to get rid of the biases completely. Whenever simulator data are used on a large scale, it is imperative to evaluate the impact of biases on these data.

7.7.4 *Comparison of real events with similar simulated situations (EDF study)*

In the literature on the subject, the number of papers is limited [Kozinsky and Pack, 1982; Beare et al., 1984; Beare et al., 1983] and show that the problem of simulator data transferability is far from solved.

In order to go further in the investigation, a study has been performed by Electricité de France with the objective of comparing real events with similar simulated situations [Dien and Peresson, 1991; Dien, 1990].

The transient chosen for a first comparison was « the spurious safety injection ».

The comparison is based on two sorts of data :

- collected objective data on the situation : time required to perform actions defined as characteristic of the situation as well as data on the environment and the situation management (times at which the shift supervisor and the safety engineer are sent and arrive, number and status of the people present in the control room...). These data are derived from the analysis of specific data sheets, for the simulator tests, and from the analysis of automatic recordings of the action chronology and from what the operators remember, for real events ;
- the operators' opinion about the differences they perceive between work on the simulator and real operation and about the impact of these differences on their behaviour. Only the operators who have actually experienced the real event are surveyed since they are in a good position to point to the differences between the two situations and compare them. The operators are interviewed soon after the occurrence of the real event.

The preliminary results of this study indicate that the method consisting in a comparison of operating experience results and tests results clarified by the tested operators' opinion seems quite appropriate to address the problem in hand.

A point stands out from the first results : with the same basic scenario, each simulated or actual situation is unique and is characterised by its own set of facts making it different from the others. However, because the chronological control actions prescribed by the procedure are scrupulously complied with, all the analyzed events belong to one common family.

We can also add that there are as many possible differences between an actual event and a simulator test as between two real events (or two simulator tests as a matter of fact). Moreover, this study demonstrates that, concerning the representativeness of the operators' behaviour, a test cannot be regarded as a whole : in other words, the results will be more or less representative of reality according to the field studied.

Thus :

- the recorded performance times give a true image of the real situation if the action performance does not depend on relations with the outside world ;
- the operators' errors are also representative because the process controlled remains identical ;

- the analysis of a diagnosis phase seems to be biased on the simulator because the simulated conditions are out of context ;
- the analysis of cognitive activities (information retrieval and processing) requires specific treatments ;
- it seems that the way an accident is perceived and lived through cannot be analyzed by observing a test.

Therefore it appears that the study performed by EDF comparing actual events and the simulation of identical events gave some interesting and positive results, since information particularly useful for PSA (like the action performance times as well as deviations) seem to be very similar during simulator tests and in real life, although it is difficult to generalise these results to all the other situations.

7.7.5 Conclusion

Operator behaviour on simulators is necessarily different to their behaviour in real-life situations. This probably has a significant effect. A list of the probable causes of differences was drawn-up. A comparison of data gathered in real-life situations and on simulators for identical situations was also carried out. The phenomena are very complex. It is not possible to establish a correlation for correcting the data from simulator tests. The « simulator effect » depends both on the events and on the parameters considered, and it can vary in opposite directions, depending on the case. It cannot be said, for example, that performance is always better on simulators. We therefore took the test data « as is » for the French PSA 900 and PSA 1300.

Simulators must not at all be rejected because of these restrictions. Operating problems observed on simulators can generally occur just as easily in real-life conditions. Simulators constitute an outstanding source of information. Their use should be encouraged for PHRAs.

7.8 Chapter References

- /1/ Hirschberg, S., Björe, S. and Jacobsson, P. (1989), "Retrospective Quantitative Analysis of Common Cause Failures and Human Interactions in Swedish PSA Studies. Proceedings of PSA'89 - International Topical Meeting on Probability, Reliability and Safety Assessment", Pittsburgh, Pennsylvania, April 2 -7, 1989.
- /2/ Swain, A.D. and H.E. Guttman, "Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications," NUREG/CR-1278, prepared by Sandia National Laboratories for US Nuclear Regulatory Commission, August 1983.
- /3/ Barriere, M., et al., *Multidisciplinary Framework for Human Reliability Analysis with an Application to Errors of Commission and Dependencies*, NUREG/CR-6265, US NRC, Washington D. C. , August, 1995.
- /4/ Beare et al. Criteria for Safety-Related Nuclear Power Plant Operator Actions : Initial Simulator to Field Data Calibration, 1983, US NRC, NUREG/CR-3092
- /5/ Beare, D. et al. A simulator-based Study of Human Errors in Nuclear Power Plant Control-Room Tasks, 1984, US NRC, NUREG/CR-3309

- /6/ Davioudian, K., Wu, J.-S. and Apostolakis, G. (1994), "The Work Process Analysis Model (WPAM): An integrated Approach to the Incorporation of Organisational Performance into Probabilistic Safety Assessment Methodology", Proceedings of PSAM-II, San Diego, CA, USA, March 20-25, 1994, pps. 080/7-13.
- /7/ Dien Y., and H. Peresson, Representativeness of the operators behaviour during tests on full scale simulators, EDF-DER Communication HT-54/91-25A, Congrès International sur la simulation, New Orleans, Avril 1991
- /8/ Dien Y., Représentativité du comportement des opérateurs lors d'essais sur simulateurs pleine échelle : cas de l'injection de sécurité surabondante - EDF-DER Memorandum HT 54/88 - 102, 1990
- /9/ Grosdeva T., and F. Mosneron-Dupin. Study of operators' activity in simulated accident situations : new orientations - Human Factors Engineering Workshop : A Task-Oriented Approach, Noordwijk, November 1989
- /10/ Grosdeva, Lemaitre, Lhuillier. Analysis of Operators' Reasoning and Reasoning-Shaping Factors, 8th European Annual Conference on Human Decision Making and Manual Control, Copenhagen, June 1989
- /11/ Hollnagel, E., "Human Reliability Analysis - Context and Control", Academic Press, NY, 1993.
- /12/ Julius, J.A., E.M. Jorgenson, G.W. Parry, and A.M. Mosleh, *A Procedure for the Analysis of Errors of Commission in a Probabilistic Safety Assessment of a Nuclear Power Plant at Full Power*, Reliability Engineering and System Safety, Vol. 50, (1995), pages 189-201.
- /13/ Julius, J.A., E.M. Jorgenson, G.W. Parry, and A.M. Mosleh, *Procedure for the Analysis of Errors of Commission During Non-Power Modes of Nuclear Power Plant Operation*, Reliability Engineering and System Safety, Vol. 53, (1996), pages 139-154.
- /14/ Kozinsky E.J. and R.W. Pack, Performance Measurement System for Training Simulators - Report EPRI NP2719-1982
- /15/ Macwan, A. and A. Mosleh, *A Methodology for Modeling Operator Errors of Commission in Probabilistic Risk Assessment*, Reliability Engineering and System Safety, 45, (1994) pp 139-157.
- /16/ Moieni, P. and Orvis, D.D. (1994), "An Approach for Incorporation of Organisational Factors into Human Reliability Analysis in PRAs", Proceedings of PSAM-II, San Diego, CA, USA, March 20-25, 1994, pps. 080/1-6.
- /17/ Parry, G.W. *Suggestions for an Improved HRA Method for Use in Probabilistic Safety Assessment*, Reliability Engineering and System Safety, Vol. 49, (1995), pages 1-12.
- /18/ Paté-Cornell, E. and Fischbeck, P. S. (1993), "PRA as a Management Tool: Organisational Factors and Risk-based Priorities for the Maintenance of the Space Shuttle Orbiter", Reliability Engineering and System Safety, Vol. 40 (1993) pps. 239-257.

- /19/ Reason, J., "Human Error," Cambridge University Press, 1990.
- /20/ Budnitz, R. J. (1997), "State-of-the-art Report on the Current Status of Methodologies for Seismic PSA. Report prepared by Future Resources Associates, Inc. for PWG5.
- /21/ Chung, G., Yee, T. and Moore, D. (1996), "SONGS 2/3 Seismic IPEEE Core Damage Risk in a High Seismic Region". Proceedings of PSA'96, International Topical Meeting on Probabilistic Safety Assessment: Moving Toward Risk-based Regulation, held in Park City, Utah, September 29 - October 3, 1996.
- /22/ Gil Montes, B. (1996), Consideration for the Development and Documentation of Human Reliability Studies for External (Fire) Events in Probabilistic Safety Analysis. Rev. 0 (June 1996), PSA and Human Factors Division, CSN, Spain.
- /23/ Oehlberg, R. N., Marteeny, M., Bateman, K., Najafi, B. and Parkinson, W. (1994), Proceedings of PSAM-II, An International Conference Devoted to the Advancements of System-based Methods for the Design and Operation of Technological Systems and Processes, held in San Diego, Ca., March 20 - 25, 1994.
- /24/ Parkinson, W., Solorzano, G., Najafi, B., Marteeny, M. and Bateman, K. (1993), "Fire Events Database for U.S. Nuclear Power Plants. EPRI NSAC-178L, Revision 1, Electric Power Research Institute, Palo Alto, Ca.

8. CURRENT DEVELOPMENT TENDENCIES

8.1 Introduction

This chapter outlines current developments in HRA methodology. These developments are intended to address some of the important outstanding issues for HRA.

These issues may be divided into “methodological” and “scope” issues. Methodological issues deal with the validation and improvement of current methods for HRA; in this regard, the replacement of current methods by methods that are better empirically or theoretically founded is not ruled out. Scope issues concern the requirements on HRA brought about by the extension of the PSA’s scope, for instance, to treat errors of commission more systematically. In contrast to methodological issues, scope issues are being addressed almost exclusively by new methods. In fact, a number of researchers are envisioning that these issues will in the long term be treated not only with new methods but with new approaches, such as dynamic simulations that integrate models of the operators and of the plant.

Section 8.2 summarises these issues, some of which were discussed in the preceding chapter. Each set of issues is presented as well as the associated research trends.

Section 8.3 reviews recently developed and emerging methods for HRA. All of the methods discussed in this section are intended to be applied in the current PSA framework.

Finally, Section 8.4 surveys the long-term prospects for HRA. Some implications of “new” applications of PSAs on the requirements on HRA are addressed in Section 8.4.1. The new applications of PSA refer, for instance, to the use of the PSA in risk-based regulation and in operational decisions. Sections 8.4.2 and 8.4.3 discuss long-term developments. These center on dynamic operator-plant models, the aims of which include both PSA (for accident sequence quantification) as well as human-machine interface design and evaluation.

8.2 Summary of Issues and Research Trends

In summarising the issues that face the HRA field, one can distinguish between methodological issues and scope issues. The first concern essentially the improvement of HRA methods whereas the latter deal with the needed extensions of the scope of analysis of human interactions with the nuclear power plant. These issues are outlined in Sections 8.2.1 and 8.2.3; Sections 8.2.2 and 8.2.4 describe some trends in research associated with these issues.

8.2.1 *Methodological issues*

Many of the current developments in the HRA field focus on issues with the currently available methods. The efforts aim to improve these methods, or in some cases, to replace the methods. This section summarises these issues. The discussion of issues concerning errors of commission and related topics are dealt with separately in Section 8.2.3. Those issues might be called “scope” issues.

Consequently, aside from not being suitable to address errors of commission, what are the problems with HRA methods? Some of the most important problems are:

- Significant differences in quantitative results from different analysts (using the same method) or from different methods
- Weak treatment of dependencies between actions. The success of a given action may be dependent on the success or failure of previous actions (in addition, of course, on the impact of the previous action on the plant response).
- Weak treatment of diagnosis, or more generally, decision-based errors.
- Lack of treatment of dependencies between PSFs (for a given action).

In discussing the causes of these problems, it should be said at the outset that the methods available today are not optimised for treating decision-based errors; their handling of errors during the execution of actions is better. The approaches to analysing decision-based errors can be divided into two major categories. One set of methods focuses on 1) whether the plant information to make the diagnosis is available, 2) the extent of training and the quality of procedures for dealing with this situation, 3) the number of concurrent tasks, and 4) the adequacy of time for making the decision.

A second set of methods focuses more exclusively on the time available; the most advanced such method is perhaps the HCR/ORE time reliability curves (TRCs). In the optimal application of these curves, the human error probability is based on the ratio of the mean time taken by operator crews in the situation being analyzed obtained in simulator studies over the time available. It can be seen that to a large degree, this approach deals implicitly with factors 1-3 (information, training/procedures, and concurrent tasks — cf. the paragraph above) since the median time to response observed in the simulator reflects all of these.

It can be seen that both sets of approaches for treating decision-based errors do not consider broadly (or not at all) the cognitive state of the operator and its evolution during the scenario. More generally, a major problem with HRA methods, speaking of the decision-related part as well as the execution part, is that their basis is relatively weak. HRA methods are not generally based on an empirically supported model of human behavior, and especially, of the cognitive behavior of operators.

To a large extent, this deficiency is caused by the lack of data on operator performance. Human performance data is difficult to collect. Human behavior (and consequently performance) is very context-specific. A large number of factors influence human error probabilities. As a result, it is difficult to compare and aggregate data from different sources (e.g. different plants) or different contexts (different scenarios). The difficulties increase when trying to use data from other industries.

From the theoretical side, psychology and cognitive science offer several models of human behavior. Some or all of these models, in particular, for problem-solving behavior, may be applicable. The problem is that their applicability will depend on the scenario, on the experience of the individuals, as well as on the specific parts of their tasks. It should be noted that uncertainty with the applicability of the different models poses problems for data collection. The human performance data collected can be useful only when the collection effort is structured with reference to a model of behavior. Currently, data still needs to be collected to generate and validate a model.

The reasons that underlie the HRA methodological problems, discussed immediately above, may be considered “intrinsic” to the HRA methods. In addition to these, some of the problems with HRA methods presented above have causes related to practical considerations. These considerations include:

The taxonomy used in HRA is not strongly Standardized. This results in discrepancies in the application of methods.

- The assumptions made by different analysts can be highly variable. Some uniformization of these assumptions may be desirable; at the minimum, it could be useful to include a standardised summary of high-level analysis assumptions. Is this a screening analysis? Are recoveries considered at the cut-set level (for important scenarios only)? Systematically, for every action?
- Proper application of some methods is in some cases regarded as too labor-intensive. One implication is that simpler methods, with strong conservatisms built-in, need to be available as alternatives to more detailed methods. Clearly, the conservatism of the simpler methods need to be validated against more detailed methods.
- Guidelines for decomposition (in methods that require it, for instance, THERP) may not be adequate. Discrepancies in the level of decomposition will generally produce different results. (Note also that a high level of decomposition is not necessarily optimal; it may lead to the failure to consider dependencies between the decomposed task elements.)

8.2.2 *Research trends related to methodological issues*

The lines of research can be divided into three main groups: basic research, development, and long-term development. Basic research in this area focuses on human behavior and human error. Development work and activities supporting development directly includes:

- HRA method development efforts
- Data collection and analysis
- Organisational factors and safety culture

Long-term development focuses on dynamic frameworks, and is intended in particular to address errors of commission and, more generally, dependencies. As mentioned earlier, this relates to the extension of the scope of PSAs. Consequently, discussion of these issues and this research appears in Sections 0 and below.

Figure 8-1 summarises the lines of research. These are discussed individually next.

<p style="text-align: center;">Basic Research</p> <p style="text-align: center;">Human behavior and human error</p> <p style="text-align: center;">Cognitive research and modeling decision diagnosis problem-solving</p> <p style="text-align: center;">Crew research</p>	<p style="text-align: center;">Development</p> <p style="text-align: center;">HRA method dev. efforts emphasize decision-based errors</p> <p style="text-align: center;">Data collection and analysis plant-specific data experience/event data</p> <p style="text-align: center;">Org. factors and safety culture</p>	<p style="text-align: center;">Long-term Development</p> <p style="text-align: center;">Errors of commission</p> <p style="text-align: center;">Dynamic dependencies (Dynamic frameworks)</p> <p style="text-align: center;"><i>discussed in 8.2.4</i></p>
---	---	---

Figure 8-1. Principal lines of research in HRA

Basic research

The main goal of basic research supporting HRA is to improve the state of knowledge concerning human behavior and human error, knowledge which is needed for the development of HRA methods. In particular, research on the cognitive behavior of operators is being conducted in a number of countries, focusing on the decision-making, diagnosis, and problem-solving processes of operators. An important component of cognitive research is cognitive modeling, in which models of cognitive behavior are being developed. Because of the dynamic nature of the problem and the need to integrate separate models for different kinds of behavior, dynamic simulations are frequently used to represent cognitive models.

In addition to the level of individual operators, research is also being conducted on crew issues, or more broadly, on group behavior. At both the individual and group levels, the issues for research are numerous enough so that it may be worthwhile to divide the work into four areas.

	Individual level	Group level
Fundamental knowledge		
Integration (Simulation)		

The cognitive simulations are intended to be used for two kinds of applications. Simulations intended for human-machine interface design and evaluation address many human factors engineering or ergonomic factors, not only problem-solving and intention formation, but also visual perception, the effects of environmental stress factors, lighting, etc. On the other hand, the simulations being developed for accident sequence simulation are more restricted in scope, focusing on the diagnosis and intention formation of the operators.

Although a few research groups have sufficient resources to carry on activities both for advancing the state of fundamental knowledge and for integrating the specific models into dynamic simulations, most groups are specialising in one or the other. Some important projects that are more oriented at advancing fundamental knowledge include the programs at the OECD Halden Reactor Project and at the Japan Atomic Energy Research Institute [Kirwan, 1994 and Tanabe, 1995]. On the other hand, examples of cognitive simulation efforts include projects in Japan, the Netherlands, Switzerland, the U.S., and the CEC Joint Research Centre at Ispra [Yoshida, 1996; Dang, 1996; Cacciabue et al., 1992].

Development

The treatment of decision-based errors is a major emphasis of recent and current efforts to develop HRA methods. Three main approaches (and combinations of these) can be identified.

- 1) Simulator-based methods. These attempt to derive HEPs more or less directly from simulator studies.
- 2) Methods based on the analysis of operational experience.
- 3) Methods based on cognitive theory.

In the area of simulator-based methods, the primary approach consists of the “decision tree” method that arose out of the HCR/ORE program. This method was originally conceived as a complement to the Time Reliability Curves [Parry et al., 1991] but is evolving independently. The decision trees represent an empirical, plant-specific ranking of performance shaping factors, based on statistics collected concerning the causal factors involved in deviations from nominal performance. Recent progress with this approach is reported in [Bareith et al., 1996].

The structure of some emerging methods is being influenced by the analysis of operational experience (operational event data). This experience or event data is identifying or highlighting performance shaping factors related to decision-making in nuclear power plant operation. In the U.S. NRC-sponsored ATHEANA project, operational event analysis is a critical element of the development of this HRA method [Cooper et al., 1996].

The third area in HRA method development consists of methods based on cognitive and psychological theory. Cognitive theory is strongly influencing the structure and taxonomy of a number of emerging methods, including ATHEANA, CREAM, and HITLINE.

The development of these methods has emphasised better foundations in cognitive theory and in general knowledge related to operator performance (in actual events and in simulator studies), with some success. Nevertheless, they can be said to “still” require a significant degree of expert judgement. Because of the strong dependence of human behavior (and hence, of human performance) on the task context, the statistics based on simulator studies are usually not in themselves adequate for quantification. Operational event data has the same weakness, being almost intrinsically anecdotal (fortunately). Nevertheless, the

consideration of both types of data in the development of HRA methods and in their application should result in improved HRAs.

Data collection and analysis has been discussed above in the context of efforts to develop HRA methods. Numerous organisations are conducting projects to collect and analyse both simulator and experience data. The use of these results in HRA method development is only one benefit of these projects. The following table summarises the uses of the data:

Table 8-1. Uses of data from plant simulators and from operational events

	Uses
Plant-specific (simulator) data collection and analysis	<ul style="list-style-type: none"> • To evaluate training effectiveness and as input to training program content • As an input to HRA methods (in their application, e.g. in “decision trees”) • To better the state of knowledge of operator behavior • As an input to HRA method development
Operational events analysis	<ul style="list-style-type: none"> • To understand the causes for events and to improve operator performance and plant safety • To better the state of knowledge of operator behavior • As an input to HRA method development

Eventually, the data may be applied to the validation of cognitive models and of HRA methods. The individual and group levels have been discussed in this discussion of issues and the related research. The organisational level is recognised to be critical to human performance and a high level of safety. Research on organisational factors and safety culture constitutes the third main area in the development-oriented research.

In this area, some efforts to develop quantification approaches that include organisational factors have been made, for instance, [Moieni and Orvis, 1994; Goldfeiz and Mosleh, 1996]. However, many efforts at this time have a more qualitative flavor. Some topics of investigation are 1) best practices and guidelines, 2) interaction of organisational forms and cultural factors. One important issue for research is that a high level of performance (which can be measured with diverse performance indicators) has not been found to be consistently correlated with specific organisational forms. A more positive trend is that some organisational traits and elements of safety culture have been identified. [e.g. IAEA, 1991]

Long-term development

Developments with long-term perspectives for HRA are discussed in Section 8.4.

8.2.3 Scope issues

In the overall framework of PSAs, two classes of human errors may be defined according to their impact on the scenario. The first, errors of omission, refers to the failure to perform an action required in response to a situation (a part of a scenario). The second class, errors of commission (EoCs), consists of the execution of actions inappropriate for the given situation. In the current state of PSA methodology and application, the comprehensive analysis of errors of commission is excluded from the scope of the studies. The extension of the scope of PSAs to include systematic treatment of EoCs is an important research objective.

Errors of commission are not exclusively decision-based errors, sometimes loosely called cognitive errors. A slip in action execution, pushing one button while intending to push another, can also result in an EoC. However, EoCs that are related to “decision” errors, that is, errors in situation assessment (diagnosis), response selection, or intention formation, tend to be more important.

First, it is generally more difficult for the operators to recognise that they have made a decision error; in HRA terms, it is more difficult to recover from an EoC due to a decision error. Simulator studies of nuclear power plant operators show higher detection rates for mistakes than for slips [e.g. Woods, 1984]. Second, the operator’s incorrect understanding of the scenario is likely to influence the performance (success probability) of subsequent required actions.

The existing PSA framework (event trees and fault trees) is ill-suited for comprehensive and systematic treatment of errors of commission. In accident sequence modeling, analysts can in fact treat errors of commission through appropriate structuring of the accident sequence model. The confusion matrix is a method for identifying scenarios in which misdiagnosis is likely, a (high-level) cause for errors of commission. In addition, human reliability analysts can identify opportunities for errors of commission in the error identification process. However, comprehensive and systematic treatment remains difficult for two main reasons.

First, the comprehensive treatment of errors of commission means that scenario branches for operator action have three outcomes (at least): performance of the required action, omission of the required action, and commission of an alternative, aggravating action. Although the ternary branch can in fact be represented in the binary success/failure structure of the existing framework, the size of the resulting event trees is significantly increased. Nevertheless, this is not inherently unsolvable.

The second and more critical difficulty is that systematic treatment of EoCs requires a complete analysis of decision errors and especially of their consequences. The “immediate” result of a decision error (an error in the decision process) is the error of commission. A second set of consequences for decision errors is their influence on the performance of subsequent required actions. Figure 8-2 uses a tree to summarise the impact of errors in decision-making, both during the situation assessment and response planning/selection phases, and in the execution phase on the overall success of a single action (task).

In this tree, A is a required action and B is some other, aggravating action. Success of A requires correct “diagnosis”, used here in the broadest sense to cover all aspects of situation assessment, correct planning or selection of a response, and execution. A slip during the execution phase may result in either an error of omission (A is omitted, for instance, due to a memory lapse) or an error of commission (the aggravating action B is committed), consequences 2 and 3 in the tree.

In branches 4, 7, and 8, the possible success of action A is not credited (the probability is assumed to be 0), in keeping with the usual PSA assumption of conservatively not crediting coincidental failures (failures that for some reason have a mitigating effect). As a result, an error in the situation assessment phase or response planning/selection phases lead only to failure (branches 9 and 10, and 5 and 6, respectively). When an incorrect response is selected either because of an error in “diagnosis” or because of an error in response planning/selection, two resulting consequences are considered. First, action A may not be part of the selected (inappropriate) response, the consequence is an error of omission (branch 5 or 9). Alternatively, the inappropriate response may call for some aggravating action B so that an error of commission occurs (branch 6 or 10).

This tree also highlights that the omission/commission classification of errors is distinct from decision-based errors (errors caused by some failure in the operator's decision processes). Decision-based errors (branches 5,6, 9, and 10) may be errors of omission as well as errors of commission. Some errors of commission are decision-based errors (branches 6 and 10), while others are not (branch 3).

For the systematic treatment of EoCs, the difficulty is not simply one of using a precise taxonomy or of adding errors of commission branches to the tree. The decision tree of Figure 8-2 is a representation of one action (A); in the current framework, its outcomes collapse into success of A, omission of A, and commission of B. The success of subsequent required actions, actions required later in a scenario, will depend not only on (are not only conditioned on) the omission of A or commission of B but on the reason for these, a slip during the execution phase, a decision error in situation assessment, or a decision error in response selection. To summarise, an adequate HRA treating decision-based errors requires the consideration of information that is not "carried" in the event tree (along each branch), information that is in fact distinct and confounded with its impacts on the system (the omission or commission).

It is worth noting that a "recovery" phase has been omitted from this tree in order to simplify the presentation. As mentioned earlier, the potential for recovering slips is higher than that for recovering decision errors. The complexity of the analysis increases significantly when recovery is considered because the recovery analysis needs to consider the expected plant response from executing the required action A or the unexpected response from committing the unrequired action B.

Thus, the main issues in expanding PSAs to handle errors of commission or, alternatively, decision errors, which are the most important types of failure leading to EoCs, are the dependencies and the need to carry along more information than is possible without expanding the tree widely. As discussed, the branches with decision errors (failures in situation assessment or response selection) have more significant impacts on subsequent actions.

To summarise, the current framework fails to carry information about decision-based errors in the tree. This information, that is, information about the full context of an action, in terms not only of the plant state but also of the evolution of the cognitive state of the operator, is needed to properly assess the potential for recovery and the success of subsequent actions. This inadequacy of the current framework is driving the research and development of dynamic approaches to address the enlargement of the scope of PSAs to treat errors of commission. This research and the prospects for tools for HRA (or for supporting HRA) is discussed in Section 8.4.

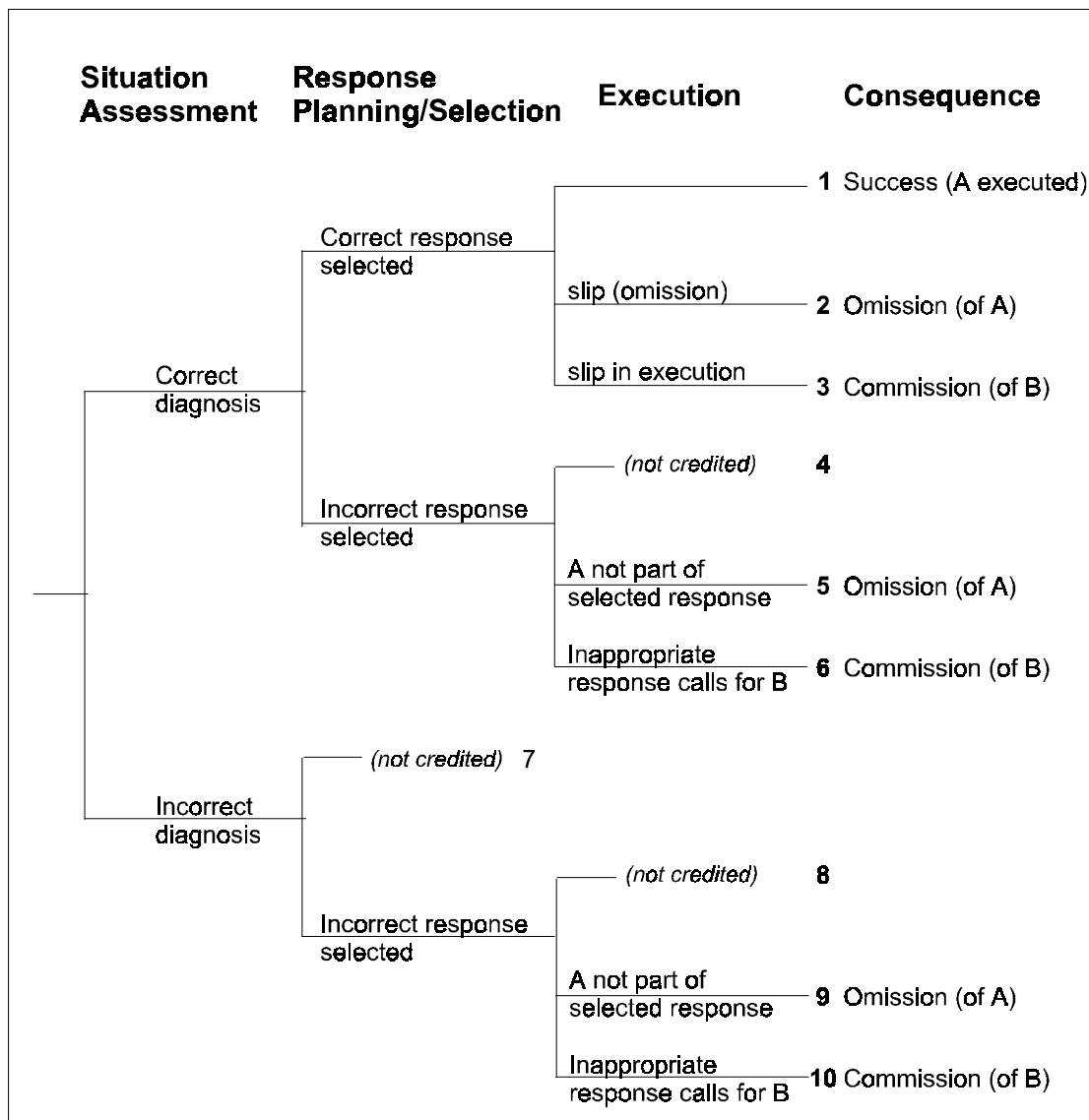


Figure 8-2. Tree representation of decision errors and their consequences.

8.2.4 Research trends related to scope issues

The enlargement of the scope of PSAs to treat errors of commission comprehensively is the motivation for two main lines of research:

- basic research to better understand human behavior and human error in the specific context of nuclear power plant operation, i.e. of operator behavior and errors
- the development of dynamic approaches for HRA or to support HRA.

As mentioned earlier, the basic research is closely related to human factors research in the area of human-machine interaction, e.g. “soft” interfaces, etc. It should be mentioned that this research could be described as quite applied. It is only “basic research” when viewed from an HRA perspective; the results of this research are relatively far away from the quantification of human actions in accident scenarios.

The dynamic approaches and the future outlook for HRA are discussed in Section 8.4.

8.3 Recently Developed and Emerging Methods

8.3.1 Overview

This section reviews recently developed and emerging methods for HRA. All of the methods discussed here are intended to be applied in the current PSA framework.

An outgrowth of the HCR/ORE methodology, the “decision tree” method based on simulator data may be viewed as representing the latest empirical quantification approach. The ATHEANA methodology, an on-going project, is an attempt to develop a new methodology from the ground up that incorporates both empirical data, mainly in the form of operational data, and psychological models of operator and human behavior. Separately, in the development of the CREAM method, the emphasis is on explicitly accounting for how context and cognition affect the cause-effect relations that underlie the failures of actions; but without modeling cognitive processing.

The screening methodology for errors of commission based on HITLINE, which has been recently applied in both a full-power PSA and a shutdown PSA, is significant as the first systematic and comprehensive methodology for identifying and screening these types of errors.

8.3.2 Simulator-based methods

HRA quantification methods that use operator performance data from simulator exercises “directly” continue to be used in some PSAs. In optimal circumstances, action-specific, scenario-specific time reliability curves (TRCs) based on simulator data specific to the plant being analyzed are preferred. In both EdF’s methodology [Berger et al, 1990] and in the HCR/ORE methodology [Moeni, 1994], TRCs of data aggregated over different required actions, scenarios, and plants are also derived and applied with some adjustments. However, for actions required in infrequent situations or actions with low failure probabilities, the validity of TRCs can be questionable and the data insufficient to derive a TRC. As a result, an approach based on “decision trees”⁸ was suggested as a complementary method to quantify HEPs. [Parry et al., 1991]

The decision trees are a plant-specific hierarchization of the main performance shaping factors (PSFs) that influence the human error probability for an action. The “*decisions*” in the tree represent the “*rating*” of the PSFs; for instance, one decision (or branching) may be “coordination”, which may have as ratings: “individual operator action” or “action requires crew coordination”. Other typical headings include “cognitive task type” (skill-based, rule-based,...), “described in procedures” (yes, no), “covered in training” (yearly). The branching points in the tree are not necessarily binary.

The HRA analyst uses the tree to represent assumptions about the importance of the different PSFs; generally, the first headings (towards the trunk of the tree) represent the more important PSFs and impact the HEP more strongly whereas the later headings (towards the leaves of the tree) represent finer differences in the HEP value. In quantification, the tree is intended to structure the evaluation of the PSFs as well as to maintain consistency in their evaluation.

8. These decision trees are not the decision trees referred to in the decision analysis field.

The tree's structure is based on the whole set of deviations observed in simulator exercises and is consequently plant-specific. The importance of the PSFs is established from an analysis of the causal factors associated with the deviations, i.e. percentage of all observed deviations caused by a given PSF. In this way, the causal factor analysis yields a global summary of the problematic factors across all scenarios and actions. It should be noted that this empirical approach is based on data about deviations between the observed responses and nominal (expected) responses to scenarios. Consequently, such deviations do not in all cases result in inadequate responses to the scenario.

Although the tree's structure is based on the deviations observed in simulator exercises, the HEP values, i.e. the HEPs associated with each set of PSF ratings (represented by a path through the tree) still have to be derived separately. In practice, these HEPs are based on other quantification methods, including THERP, and expert judgement. Each path in the decision tree thus represents an action with a given set of generic PSF characteristics, which is quantified separately. The graphical representation offered by the tree helps to visualise the relative impact of the PSFs on the HEPs as well as to identify the most effective potential error-reduction measures.

Comprehensive applications of the decision tree, with correspondingly large-scale data collection programs, include the PSA of the Hungarian Paks nuclear power plant [Bareith et al, 1996]. The use of decision trees allows qualitative data about causal factors to be integrated into the HRA, in addition to the quantitative time response data. The data collection effort and the causal factor analysis can generate useful insights for training.

The HRMS and the related JHEDI computerised methods are analogous to these decision trees in that extrapolation is performed data points based on performance shaping factors, as described in [Kirwan, 1994]. This extrapolation is based on expert judgement but is based on actual data. However, the data are proprietary and hence the methods are not publicly available. These methods have shown good results [Kirwan et al, 1994] although they are limited to skill- and rule- based actions because of the lack of data for cognitively challenging tasks.

8.3.3 ATHEANA

The ATHEANA project is an emerging HRA methodology intended to address both errors of omission and errors of commission within the existing PSA framework. The fundamental premise of the methodology is that important post-initiating event human failure events, especially errors of commission, represent situations in which the plant conditions and performance shaping factors (PSFs) virtually force operators to fail. A brief overview of the methodology is presented in this section; the process is described in more detail in Appendix E. In this overview, the terminology introduced in ATHEANA is used.

Besides a new quantification model, ATHEANA (A Technique for Human Error Analysis) suggests a new approach to human error identification. The overall methodology consists of five tasks:

Task 1.	Familiarisation with PRA model and accident scenarios
Task 2.	Identification of potential <i>human failure events</i> (HFEs) and associated <i>unsafe actions</i> (UAs)
Task 3.	Identification of the most probable/significant causes of the <i>unsafe actions</i> (the <i>error forcing contexts</i> or EFCs)
Task 4.	Refinement of HFE definitions and integration into PRA logic model
Task 5.	Estimate the likelihoods of the <i>error forcing contexts</i> and the consequential probabilities of unsafe actions

These tasks are described next along with definitions of the terminology.

Task 1 is the basic starting point for human reliability analyses in general. PRA models define the situations in which operator actions need to be performed. (In task 4, which will be discussed shortly, the PRA models can be modified to account for findings resulting from the HRA, for instance, that the operators are likely to take some action not accounted for in the PRA models.)

In task 2, *human failure events* are identified. *Human failure events* (HFEs) are human-caused failures at the function, system, or component level; they have similar impacts on the system as hardware failures. The HFE may result from either operator errors of omission or errors of commission. The potential *unsafe actions* (UAs) associated with each HFE are then identified. *Unsafe actions* are those actions taken, or not taken when needed, by plant personnel, that produce the HFE.

Task 3 is the methodological step in which ATHEANA differs most significantly from other HRA quantification approaches. In this task, the analysts examine the situation to identify potential factors that make unsafe action (UA) likely, *force* the UA. Each set of factors is referred to as an *error forcing context* (EFC). The search considers

- 1) formal (proceduralized) and informal (resulting from training) rules that would motivate the UA
- 2) ways in which the situation would appear to match the rules that call for the UA
- 3) potential ways in which the operators could recognise the inappropriateness of their mental model of the situation

In task 4, the PRA accident model is modified to include the identified *human failure events* (HFEs).

Finally, in task 5, the probabilities of the HFEs are quantified by estimating the probabilities of the error forcing contexts (EFCs) and the conditional probabilities of the unsafe actions (UAs) given each EFC.

- The estimation of the EFC probabilities is intended to be based on models of the joint probabilities of the elements of the EFC. This has been shown to be feasible in the trial application of ATHEANA. [Cooper et al, 1996.]
- A method for quantifying the conditional probabilities of the UAs still needs to be developed. Nevertheless, it should be noted that in some cases, the EFC will “almost guarantee” the UA.

8.3.4 CREAM

The development of the Cognitive Reliability and Error Analysis Method (CREAM) emphasises the importance of the context in determining human performance and the intrinsic role of cognition in all actions, and hence in all errors. CREAM attempts to explicitly account for how context and cognition affect the cause-effect relations that underlie the failures of actions. For CREAM, this explicit account (or model) is the Cognitive Control Model (COCOM); it provides the basis for associating causes and effects in the realm of cognition.

This summary description of CREAM is based on [Hollnagel, 1994].

Causes and effects are classified based on distinguishing between *error modes or manifestations (phenotypes)* and their *causes (genotypes)*. *Error modes* or phenotypes include, for instance, “action at wrong time” and “action on wrong object”; these error modes are further subdivided. The *error causes* fall into twelve groups; some examples of these are causes associated with *interpretation, procedures, the interface, and communication*. The twelve groups can be divided more generally into person-related, system-related, and environment-related.

Each group of error causes is itself described in terms of relations of causes and effects. That is, each cause in the twelve groups has potential causes; some causes for *communication failures* (as a cause of the error mode “action at wrong time”) include *distraction, inattention, etc.* In effect, a network between causes and effects is established; the network structure of this classification scheme allows a more flexible analysis. It can thus be seen that the set of possible causes for failure is very large; for HRA, the propagation through the network of causes needs to be constrained.

The CREAM method relies on a systematic description of the context to narrow down what causes and error modes should be considered. In fact, the analysis begins with an analysis of the application (a task analysis). In traditional HRA methods, performance shaping factors (PSFs) are determined for a given action. In contrast, in applying the CREAM method, the *common performance conditions (CPCs)*, which apply to the performance as a whole, are described first, before considering actions.

The steps in a CREAM analysis are:

1. Application analysis (task analysis)
2. Context description
3. Specification of target events
4. Qualitative performance analysis
5. Selection of events for further analysis
6. Quantitative performance prediction

The application or task analysis (step 1) should consider not only the operator and control tasks but also the organization and the technical system. In step 2, the context is described in terms of the CPCs. This set of conditions will determine the control mode, the way in which actions are chosen and carried out, and thus, potential errors causes.

The specification of the target events (step 3), the human failure events to be analyzed, is based on the results of both PSA analyses and the task analysis. As in the general HRA framework, the PSA analysis provides a set of required actions; the task analysis determines possible *error modes*, which correspond to the human failure events.

In step 4, the context description in terms of CPCs is used to describe (enumerate) possible causes for a target event. For each target event, both the general groups of causes and then specific causes are identified in view of the context; that is, the qualitative performance analysis identifies potential and likely causes based on the context. The relationships between the conditions for the performance (CPCs)

and error causes are described by the COCOM model, which describes the mode of operator behavior (the control mode) and the associated error modes under different performance conditions.

In PSA, a quantitative prediction of performance is usually needed. As a result, most events will be selected for further analysis (quantification). In step 6, the human error probability, i.e. the probability for the target event, is derived by assigning probabilities to the error modes given the CPCs. For now, expert judgement is necessary. The expert judgement process, however, is supported by the description of the context. The long-term aim would be to collect empirical data to supplement, calibrate, or replace expert judgement.

Although expert judgement is unavoidable in CREAM, the method appears promising because the cause-effect classification scheme provides guidance for data collection and analysis. The assignment of probabilities for the error modes given the CPCs could be eventually supported by statistical analysis of actions and their contexts (the CPCs).

8.3.5 *HITLINE*

The HITLINE methodology is the basis for a methodology for the screening of errors of commission, which is described in Chapter 7 of this report. (The methodology and results of this screening analysis are also described in the Dutch questionnaire response submitted in connection with this task of PWG5.)

HITLINE is a methodology based on simulation for treating errors of commission in PSAs. [Macwan and Mosleh, 1994] The HITLINE tool systematically generates sequences of operator actions within a dynamic event tree by considering a set of comprehensive performance influencing factors (PIFs) at each branching point. Sections 8.4.3 and 8.4.4 describe dynamic event trees (DETs) in more detail; for further information about DETs, refer to [Siu, 1992].

The HITLINE model consists of several sets of mapping rules, which

- generate and update the values of PIFs based on current plant information and the position in the emergency operating procedures
- describe the influence of the PIFs on the probabilities of the operator actions

The mappings have been developed based on the literature, operating events, as well as expert judgement.

The PIFs are discrete-valued and cover

- crew characteristics, e.g. level of training and experience,
- plant-related information such as the values of critical parameters,
- factors related to the procedures, e.g. the logic structure and number of logical conditions, and
- operator-related factors, e.g. the current diagnosis, and expectations of plant response.

The operator actions of interest are defined in relation to the response required in the procedures. At the global level, they are *incorrect procedure selection* and *delayed procedure selection*; at the local level,

they are *skip*, *delayed action*, and *commission* (of an action different from that prescribed by the procedure).

By representing the influence of PIFs on the operator response as mappings, the proposed methodology is systematic and traceable. Further work needs to be carried out to validate the mapping rules, including the associated weights (conditional probabilities). The methodology is in development and has been demonstrated through a hypothetical example described in [Macwan and Mosleh, 1994].

8.4 Prospective Outlook For HRA

8.4.1 HRA in Light of New PSA Applications

This section highlights some issues for HRA related to recent applications of PSA. Two of these uses of PSAs are discussed:

- risk-based regulation
- applications to support operations

Some examples of risk-based regulation applications include licensing, modifications to technical specifications, and exemptions from plant modification requirements based on probabilistic arguments (PSA results). In operations support, the examples relate to the concept of “risk monitors”, models that predict the changes in risk in different plant states (configuration management). Outage planning that considers the impact on risk quantitatively is a second example in this area.

The use of PSA results in regulatory decision-making raises two issues for the performance of PSA studies.⁹ First, acceptable methods and modeling assumptions for performing a PSA study have to be defined. Second, comparability between analyses becomes critical; in particular, the concern lies with the potential for decisions that differ because of differences between methods and assumptions rather than because of actual differences between plants. These issues apply to all types of analyses used in a PSA but may be particularly important for HRA.

To address both the acceptability of analyses and comparability, standardisation has been suggested. Based on a concise summary of the state of HRA, [Parry, 1996] suggests that standardisation first needs to address the issues that should be treated by HRA methods, as well as the factors that need to be considered in the HRA. At this time, the state of HRA methodology precludes agreement on one or more “standard” methods. Furthermore, Parry points out that it is insufficient to focus on quantification models; the assumptions that underlie the PSA models have a critical impact on the HRA. In summary, acceptability and comparability hinge on agreement on:

- the issues to be treated by HRA methods, rather than on acceptable methods
- an adequate level of documentation, both of quantification assumptions and of modeling assumptions

9. Direct policy issues are not discussed here, for instance, whether safety goals should be applied to individual plants or to the population of plants. A discussion of this and other issues may be found in [El-Bassioni et al., 1996].

In the process of reaching a consensus on these issues, there is also a need to recognise that the HRA discipline is currently in a period of broad changes, especially as approaches are proposed to handle the unresolved issue of errors of commission.

For applications of PSA in supporting operations, a quick requantification of models under different sets of conditions needs to be possible. For instance, one survey suggests that outage management groups would like feedback from the risk model within about 20 minutes of a change in the outage work schedule. [Hewitt and Rao, 1996]

For HRA, one set of difficulties caused by this requirement is related to the quantification assumptions. The validity of the assumptions of the HRA quantification needs to be assessed relatively quickly. For on-line risk monitors, an automatic process for checking these assumptions and indicating when the model is outside the range of validity would be needed. Even for other applications where PSA specialists can change the model, requantification of the HEPs needs to be possible. It can be seen that as for regulatory applications, the application of the PSA to consider different plant configurations also raises the requirements on the level of documentation.

8.4.2 Issues in the long term

The issues for HRA in the longer term are:

- the collection of data for validating methods
- the collection of data as input to HRA methods
- knowledge about operator behaviors, especially in cognitively challenging scenarios

The lack of human performance data and the programs for collecting and analysing this data have been discussed earlier in this chapter. It is hoped that the collection of data will be made more straightforward as a result of the empirical and theoretical bases for methods such as ATHEANA and CREAM.

For dynamic operator-plant models, which are currently viewed as eventual solutions for addressing errors of commission, additional knowledge about operator behaviors is still needed. The prospects for these models are discussed further in the following section.

8.4.3 Dynamic operator-plant models

In the context of HRA and PSA, dynamic simulations of operators and plant are being developed as tools either to quantify accident sequences or to support the HRA analysis. (Other operator-plant simulations are also being developed for human factors applications, for instance, to evaluate proposed human-machine interface designs. CAMEO is an example of such an effort [Fujita et al, 1993].)

Although different frameworks are being proposed, operator-plant simulations consist of coupled **interacting models** of the plant and of the operators. In this way, the interdependent behaviors of plant and operators can be addressed explicitly. **To support the HRA analysis**, especially the analysis of errors of commission, simulation tools can be used to:

- identify potential situations where the judgement of the operators concerning the appropriate response is inconsistent with the procedures

- study the consequences of errors of commission and the possibilities for recovering from such errors

The general idea is to simulate the operator response with a cognitive model, which addresses the evolution of the operators' mental state as they follow procedural guidance and their training. The cognitive model treats the information processing behavior of the operator.

It is important to identify situations where the operators' assessment of the dynamically evolving situation and of the appropriate response may be inconsistent with the procedures; such mismatches could potentially drive incorrect interpretation of the procedures as well as deviations from the procedures. In any case, the conflict of the response required by procedures and training with the operators' own judgements will certainly increase stress and hinder performance.

The evolution of the operators' mental state is also important to correctly model the dependencies between required operator actions. For example, when an incorrect diagnosis underlies an error of commission, this inappropriate diagnosis is also likely to influence the performance of subsequent actions required in a scenario. In addition, whether errors (of omission or of commission) are due to mistakes or slips also affects the probability of recovery. Thus, operator-plant models are being developed to treat not only the consequences of errors of commission, that is, the impact of the error on the plant response, but also the consequences of the underlying, cognition- or decision-related failures on subsequent operator performance.

Table 8-2 lists some dynamic operator-plant models in development. In the top half of the table are models whose characteristics suggest that they are directed primarily towards supporting HRA and accident sequence analyses (supporting the calculation or estimation of HEPs); the remaining models are aimed at accident sequence quantification, which are discussed next.

The operator-plant simulation models developed for **accident sequence quantification** have been influenced by the requirement that the "space" of possible sequences has to be adequately covered. The models aimed at sequence quantification have usually been expressed in a dynamic event tree framework. In contrast to analog simulations, where stochastic sequences (sequences involving probabilistic branching events) are simulated sequence-by-sequence by sampling each branching event, dynamic event trees simulate all sequences concurrently, without sampling the branching events.

Table 8-2. Dynamic operator-plant models for HRA-related applications

Operator-plant models in development	References
For supporting HRA and accident sequence analysis	
CES	[Woods et al., 1989]
COCOM	[Hollnagel, 1993]
COSIMO	[Cacciabue et al., 1992]
JACOS	[Yoshida et al., 1996]
OPSIM	[Dang, 1996]
For direct accident sequence quantification	
ADS	[Hsueh et al., 1994]
HERMES	[Cacciabue et al., 1996]
HITLINE	[Macwan and Mosleh, 1994]

These alternative dynamic frameworks are discussed further in the next section. A detailed comparison of dynamic event trees and analog simulations is presented in [Siu, 1992]. The dynamic event tree framework influences the operator model in particular in two ways:

- the state of the operator model (the operator state) has to be explicitly expressible
- the operator's behavior (and cognitive states) tends to be modeled at a coarser level of granularity

The operator state has to be explicitly expressible so that it can be processed and then stored, as the simulation "grows" each branch of the dynamic event tree at each time step. Besides keeping the dimensionality of the operator state low, the coarser-grained structure of the models reflect the focus on the connections between the actions taken by the operators and the underlying mental states.

In other words, cognitive processing is not explicitly modeled, as is more typical in analog simulations. For instance, rather than modeling the processing of each item of information to derive the operator's diagnosis and intentions, the diagnosis process is modeled as probabilistic transitions between diagnosis states (reaching different situation assessments).

Both types of operator-plant models require data about operator behavior for further development and eventually for validation. For the cognitive models, which are generally intended as supporting tools for performing HRA and PSA, information about the cognitive behaviors of operators is needed. One of the first priorities is to establish the cognitive processes and their characteristics that need to be included. Consequently, the required data tend to be more (but not exclusively) qualitative; quantitative data would also be needed to test model hypotheses.

For quantification-oriented models, data is needed to obtain the transition probabilities for input to these models. The transition probabilities would be derived from or estimated. It should be noted that these probabilities are not analogous to human error probabilities in traditional HRA; instead, they represent probabilities in the evolution of the operator mental state. Some of the bases for transition probabilities, e.g. whether the operators select the correct response based on the current diagnosis state, would include the level of cognitive processing expected in this situation.

8.4.4 *Simulation-based frameworks for PSA*

An adequate treatment of the dependencies associated with errors of commission, of which decision-based errors are of particular concern, seems to require dynamic analysis (simulation) tools to treat the evolution of the operators' mental state. Some operator-plant models discussed in the previous section are intended primarily for such analyses, the results of which would be used to structure the accident sequence model and to assess human error probabilities in conventional analyses.

It may be argued that the state of knowledge concerning operator cognitive behavior and the difficulty in assessing probabilities for specific behaviors will remain inadequate for such models to ever be quantitative. A level of quantification would be achievable by describing what is more likely and less likely and by express such likelihoods as probabilities. Probabilities may represent both real uncertainty, i.e. whether an operator will behave in this or that manner, as well as modeling uncertainty, i.e. which behavior model is applicable.

In any case, the static or semi-static representation of the conventional model may not be an optimal approach to what appears to be inherently a dynamic problem. Consequently, a number of approaches

have been proposed based on simulation of some type. These methodologies or frameworks have been labeled “revolutionary”, since they would replace the fault tree/event tree framework (at least for the analysis of some scenarios).

The type of simulation of choice for revolutionary approaches has been the dynamic event tree (DET): DYLAM [Amendola, 1984], DETAM [Acosta and Siu, 1991], ADS [Hsueh et al, 1994]. Since they do not sample probabilities, DETs ensure coverage of the probability space. However, the dynamic event tree expands dramatically and requires algorithms for dynamically limiting the size of the tree, that is, the simulation of some branches should be stopped on the basis of a probability cut-off (the branch becomes too improbable to be a significant contributor) or of other rule. The merging of branches with similar characteristics is also an alternative. A second issue for dynamic event trees is that they require a relatively simple model of the operator. Operator models that are inherently simple, such as COCOM [Hollnagel, 1993], or empirical models would be preferred. Other models could be developed by simplifying a more complex model (much like a response surface or linear model can be used in place of a phenomenological model). However, inherently simple models may be more promising since the more detailed, complex models of operators are still in development.

The issues faced by analog simulations, the main alternative, are complementary. As a depth-first approach, the models can in principle be arbitrarily complex. However, the “rare event” problem is faced in the sampling of the sequences. In other words, the concern is that rare, problematic sequences may not be sampled. Some sampling methods, such as importance sampling, can be used to address this problem.

For HRA, in conclusion, analog simulations will tend to be oriented to supporting analyses, in which detailed models are necessary. They allow flexibility in the design of the operator model. Dynamic event trees are favored for applications where direct accident sequence quantification is the aim. Their disadvantage, in turn, lie in the requirement for simplified models.

8.5 Chapter references

- /1/ Acosta, C.G. and N. Siu, “Dynamic Event Tree Analysis Method (DETAM) for Accident Sequence Analysis, MITNE-294, Massachusetts Institute of Technology, 1991.
- /2/ Amendola, A. and G. Reina, “DYLAM-1: A Software Package for Event Sequences and Consequences Spectrum Methodology,” EUR-9224 EN, CEC-JRC, 1984.
- /3/ Bareith, A., , E. Holló, S. Borbély, A.J. Spurgin, “Treatment of Human Factors for Safety Improvements at the Paks Nuclear Power Plant.” In P.C. Cacciabue, I.A. Papazoglou (Eds.), *Proceedings of ESREL’96 - PSAM III International Conference on Probabilistic Safety Assessment and Management*. Crete, Greece, 24-28 June, 1996, pp. 1191-1196. Springer-Verlag London, UK.
- /4/ Berger, J.-P. et al., “EPS-1300 - Rapport de Synthèse,” Electricité de France (1990) English translation available.
- /5/ Cacciabue P.C., Cojazzi, G., Parisi, P. (1996). “Dynamic HRA Method Based on a Taxonomy and a Cognitive Simulation Model.” In P.C. Cacciabue, I.A. Papazoglou (Eds.) *Proceedings of ESREL’96 - PSAM III International Conference on Probabilistic Safety Assessment and Management*. Crete, Greece, 24-28 June, 1996.

- /6/ Cacciabue, P.C., F. Decortis, B. Drozdowicz, M. Masson, J.P. Nordvik, "COSIMO: a cognitive simulation model of human decision making and behavior in accident management of complex plants," *IEEE Trans. Sys., Man, and Cybernetics*, Vol. 22 (1992), pps. 1058-1074.
- /7/ Cooper, S.E., A. M. Ramey-Smith, J. Wreathall, G. W. Parry, D.C. Bley, W.J. Luckas, J.H. Luckas, J.H. Taylor, M.T. Barriere, "A Technique for Human Error Analysis (ATHEANA), Technical Basis and Methodology Description," NUREG/CR-6350, U.S. NRC, Washington, DC, 1996.
- /8/ Cooper, S.E., Ramey-Smith, A., Wreathall, J., Parry, G.W., Bley, D.E., Taylor, J.H., and Luckas, W.J., *A Technique for Human Error Analysis (ATHEANA) - Technical Basis and Methodology Description*, DRAFT NUREG/CR-6350, 1996.
- /9/ Dang, V.N., "Modeling operator cognition for accident sequence analysis: Development of an operator-plant simulation". Ph.D. thesis. Department of Nuclear Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. May 1996.
- /10/ El-Bassioni, A., J. Guttman, W.B. Hardin, A. Ramey-Smith, H. Woods, M. Cunningham, W.T. Pratt, A. Camp, "Issues associated with risk-informed regulation of nuclear power plants" In P.C. Cacciabue, I.A. Papazoglou (Eds.) *Proceedings of ESREL'96 - PSAM III International Conference on Probabilistic Safety Assessment and Management*. pp. 738-741. Crete, Greece, 24-28 June, 1996.
- /11/ Fujita, Y., I. Yanagisawa, K. Nakata, J. Itoh, N. Yamane, R. Kubota, and M. Tani, "Modeling Operator with Task Analysis in Mind", *Proceedings of the ANS Topical Meeting on Nuclear Power Plant Instrumentation, Control, and Man-Machine Interface Technologies*, Knoxville, TN, USA, April 18-21, 1993, pp. 505-512.
- /12/ G. Parry, "Standardizing Human Reliability Analysis – Issues and Suggestions," *Proceedings of PSA '96 International Topical Meeting on Probabilistic Safety Assessment*, Park City, UT, US, 29 Sep. - 3 Oct., 1996, pp. 1243-1247.
- /13/ Goldfeiz, E.B., A. Mosleh, "A Methodology for Explicit Inclusion of Organisational Factors in Probabilistic Safety Assessment," In P.C. Cacciabue, I.A. Papazoglou (Eds.), *Proceedings of ESREL'96 - PSAM III International Conference on Probabilistic Safety Assessment and Management*. Crete, Greece, 24-28 June, 1996, pp. 916-921. Springer-Verlag London, UK.
- /14/ Hewitt, J. and D. Rao, "Enhancing PSA Use in Outage Scheduling Applications at Grand Gulf Nuclear Station", *Proceedings of PSA '96 International Topical Meeting on Probabilistic Safety Assessment*, Park City, UT, US, 29 Sep. - 3 Oct., 1996, pp. 597-599.
- /15/ Hollnagel, E., *Principles of Cognitive Reliability Analysis*, based on a presentation given at the International Workshop on "Human Reliability Models: Theoretical and Practical Challenges" (Series on Advanced Topics in Reliability and Risk Analysis), Stockholm, August 22-24, 1994.
- /16/ Hollnagel, E., *Human Reliability Analysis: Context and Control*. Academic Press, London, 1993.

- /17/ Hsueh, K.-S., L. Soth, and A. Mosleh, "A Simulation Study of Errors of Commission in Nuclear Accidents," Proceedings of PSAM-II, San Diego, CA, USA, March 20-25, 1994, pps. 066/1-6.
- /18/ International Nuclear Safety Advisory Group, "Safety culture," Safety Series No. 75-INSAG-4, International Atomic Energy Agency, 1991.
- /19/ Kirwan, B., "Human Error Project Experimental Programme," HWR-378, OECD Halden Reactor Project, Halden, Norway, November 1994.
- /20/ Kirwan, B., *A Guide to Practical Human Reliability Assessment*, Taylor and Francis, London, 1994.
- /21/ Kirwan, B., Kennedy, R., Taylor-Adams, S. and Lambert, B. (1994), "Validation of Three Human Reliability Assessment Techniques: THERP, HEART and JHEDI." Volume I - Main Report, University of Birmingham, February 1994.
- /22/ Macwan, A. and A. Mosleh, "A methodology for modeling operator errors of commission in probabilistic risk assessment," *Reliability Engineering and System Safety*, Vol. 45, pp. 139-157, 1994.
- /23/ Moieni, P., A.J. Spurgin, and A. Singh, "Advances in Human Reliability Analysis Methodology. Part I: Frameworks, Models and Data. Part II. PC-based HRA Software," *Reliability Engineering and System Safety*, Vol. 44, pp. 27-66, 1994.
- /24/ Moieni, P., and Orvis, D.D., "An Approach for Incorporation of Organisational Factors into Human Reliability Analysis in PRAs", Proceedings of PSAM-II, San Diego, CA, USA, March 20-25, 1994, pps. 080/1-6.
- /25/ Parry, G. W., Singh, A., Spurgin, A., Moieni, P. and Beare, A. (1991), An Approach to the Analysis of Operating Crew Response Using Simulator Exercises for Use in PRAs. OECD/BMV Workshop on Special Issues of Level 1 PSA, Cologne, Germany, 28 May 1991.
- /26/ Siu, N., "Risk Assessment for Dynamic Systems: An Overview," *Reliability Engineering and System Safety*, submitted for publication, Apr. 1992.
- /27/ Tanabe, F., "Overview of Human Factors Research in JAERI", International Conference on Probabilistic Safety Assessment, Seoul, Korea, Nov. 26-30, 1995.
- /28/ Woods, D.D., "Some results on operator performance in emergency events," Institute of Chemical Engineers Symposium Series, 1984, 90, pps. 21-31. (In *Ergonomics Problems in Process Operations*, Birmingham, U.K., 11-13 July 1984.)
- /29/ Woods, D.D., H.E. Pople, and E.M. Roth. "The Cognitive Environment Simulation as a Tool for Modeling Human Performance and Reliability," NUREG/CR-5213, U.S. Nuclear Regulatory Commission, Washington, DC, 1989.
- /30/ Yoshida, K., Yokobayashi, M., Tanabe, F., and Kawase, K., "Development of AI-Based Simulation System for Man-Machine System Behavior in Accidental Situations of Nuclear Power Plant," *J. of Nuclear Science and Technology*, Vol. 33, No. 2, 1996.

9. CONCLUSIONS AND RECOMMENDATIONS

The present state-of-the-art report provides an overview of the HRA-related research programs and applications in Member countries, reviews HRA methods currently used in these countries based on a large selection of representative PSAs, surveys HRA practices and their results, and outlines trends in methodological developments.

Table 9-1 displays the specific areas of research by country. The picture is certainly not complete since the emphasis in the summaries of research provided by the task contributors has been on HRA. For example, activities on “Organisational factors” have been included only when their objectives are related to some extent to PSA activities. Obviously, most of the countries have activities in the field of organisation and management when regarded from a broader human performance perspective.

Table 9-1. Areas of research related to human performance listed by country

	HRA methods development	Plant-specific data collection	Operational event analysis	Maintenance - analysis / support tools	Human behavior and human error research	Crew issues	Simulation of human-machine system	Organizational factors	Operator aids	Computerized procedures	Advanced (soft) interfaces	Accident management
B								•		•		
Can	•	•										
CH	•		•		•							
Cz		•										
Fin	•	•	•	•								
Fra		•							•	•		
Ger	•		•		•		•	•				•
Ita			•		•			•	•	•		
Jap	•	•	•		•	•	•	•	•	•		
Kor	•		•		•	•						•
NL	•		•		•						•	
UK	•	•	•									
US	•	•	•				•	•	•	•		
EU	•				•	•					•	

From: CSNI PWG 5 Task 94-1
Summaries of activities in HRA/HF

As can be seen, most of the countries participating in the task have active programs in “HRA methods development” and “Operational event analysis”. It should be noted that points in the table only indicate the existence of specific activities; the resources allocated to these projects vary significantly, from very small to fairly large.

The survey of international HRA practices has demonstrated that the uses of standard PSA techniques have significantly matured in the last few years. This may be attributed to the establishment of accepted approaches, increased experience in applying them, and to the impact of more extensive and demanding review procedures. Nevertheless, there are significant differences between the studies in terms of the implementation of the methods used. This aspect was evident when some important operator actions, characteristic for BWRs and PWRs, were subjected to a detailed treatment. While there was no intention to directly compare the HEPs since the plants represent a variety of designs and since differences exist with respect to the degree of automation, procedures and context of the actions, the factors that drive the numerical values could be identified.

A wide spectrum of HRA methods were represented in the responses provided by the participating countries. This includes: (a) Decomposition or Database Techniques (THERP/ASEP, HEART); (b) Time Dependent Methods (OAT/TRCs, HCR, HCR/ORE); (c) Expert Judgement Based Techniques (APJ; SLIM/FLIM).

Only a few studies use a single method; in most cases several techniques were combined. The responses provide a number of insights concerning mixing and matching different methods, criteria in the choice of HRA overall approach, and views of reviewers and PSA users.

Current developments in HRA methodology focus on dynamic actions and are motivated by four weaknesses of existing methods. **First**, existing methods rely to a large extent on expert judgement, due to scarcity of empirical data on human performance. **Second**, the current methodology does not adequately identify, explicitly represent and quantify the likelihood of actions with potential adverse effects on plant conditions (errors of commission). **Third**, explicit accounting for dependencies among actions is left to the event tree/fault tree structures that represent the human actions. Consequently, these methods have a great deal of room for improvement in modelling dependencies. **Fourth**, it is known that the organisational culture and its safety-related aspects known as safety culture has an important impact on risk. On the other hand, quantification of this impact or more generally, the impact of management and organisational factors, remains an open methodological issue for PSA/QRA. Generally, current methods provide only a means to quantify failures based upon broad categories of failure “modes”, and do not address the underlying causes and mechanisms of failure.

In spite of these serious limitations of currently used methods the HRA survey conducted in the present task demonstrates that PSAs have been successful in terms of identifying deficiencies related to human performance. Thus, more than 40 specific examples of HRA-related improvements of hardware and procedures were described in the responses based on 21 PSAs included in the survey. The modifications concern: new procedures, revision of procedures or technical specifications, installation of “new” systems or automated capabilities, and modification of systems (including actuation logic).

The work in progress on HRA methodology can be divided into three categories. The first group consists primarily of efforts to validate or calibrate existing methods and to develop methods on the basis of empirical data. Because these efforts are closely related to existing methods and are intended for use within the current PSA framework, the empirically-based methods have been called evolutionary methods. ATHEANA, which explicitly attempts to address errors of commission, belongs to this category.

In the second group of developments is longer-term research to treat the commission aspect of operator failures taking into account the evolution of the cognitive state of the operators in a dynamic framework. Since the scope of the HRA would be extended and the PSA framework may require significant modifications to allow dynamics to be modelled more explicitly, the methods in development are often characterised as revolutionary. Some researchers have suggested that these methods should be based on a model of human performance supported by data specific to the task domain. Consequently, for this second group of efforts and methods, empirical data is needed, first, to develop and validate the model of human performance, secondly, for the development (and validation) of the HRA methodology itself, and, finally, as an input when using the HRA methods.

The third group includes full scope cognitive models intending to apply a more fundamental understanding of the causes and mechanisms of operator response, they tend to be very complex and typically require very detailed knowledge and representation of the context (e.g. the accident environment including system behaviour, procedural instructions, crew interactions, and accident history). Those that have reached the implementation level are either purely or partially simulation-based, and seem to require a different PSA environment. The actual or anticipated complexity of these models has resulted in a rather sceptical and cold reaction among many traditional HRA analysts and potential users. The legitimate concern for practicality is, however, sometimes mixed with the point of view that we do not know enough about human behaviour and that the more complex models are merely based on conjectures and do not necessarily improve the quality of our predictions. In spite of these doubts there is a significant momentum towards developing a better model and there are signs of progress.

Approaches to the systematic incorporation of organisational factors have been recently proposed. As with performance shaping factors, there is a consensus on factors that affect safety negatively and at the same time inadequate knowledge about the interactions between organisational factors. Methodologies are still being developed.

Based on the present study the following recommendations can be made:

1. There have been surprisingly few attempts to validate the HRA methods that are currently widely used. While the limitation of these methods to treat errors of commission is inherent, it is likely that in the medium-term perspective these methods will continue to be predominantly used in industrial PSAs. For this reason validations within their scope of applicability should be encouraged.
2. As demonstrated by the survey carried out within the present task, in spite of increased maturity of PSAs there are quite large differences in the implementations of standard methods. The current trend towards risk-informed regulation calls for improved homogeneity in the ways the methods are applied. Some researchers, practitioners and reviewers propose quite strict standardisation of HRA. This should not be driven to such extent that the incentive for further developments decreases, resulting in stagnation.
3. Data collection and analysis efforts need to be intensified. This includes operational events studies that are important as an input to method development. The use of simulators is extremely beneficial not only in the context of training but also for generating data for methods (analyses), and for methods evaluation/validation. Regrettably, today there are very few active simulator programmes supporting HRA. Hopefully, the few existing experimental programmes researching Performance Shaping Factors (PSFs) will be in the future more

directly connected to HRA methods development and will explicitly consider PSA experiences and perspectives when designing the experiments.

4. The various emerging HRA methods are at different stages of development. It is still an open question to what extent they will fulfil the originally set objectives in the transition from concepts through case studies to full scope implementations and applications. There is a good reason to closely follow these developments.
5. The current PSA framework with its schematic representation of sequences can only to a limited extent and primarily through improvisation support systematic and comprehensive treatment of cognition-dependent operator errors. This inadequacy is the reason why development of dynamic approaches, based on different forms of simulation, is of high interest. In the intermediate term dynamic applications with more restricted objectives could be very useful. Such applications include dynamic verification and safety assessment of procedures, both standard written or computerised, or of operator aids.
6. In view of the need to develop (in relative terms) more human-centred PSAs, the need of multi-disciplinary competence and the relatively small resources being available for this task, the importance of co-operations cannot be overemphasised. Such co-operations should involve both human reliability analysts and human factors specialists; in particular, there is a good but far from fully used potential for international co-operations in this context.

APPENDIX A. DESCRIPTION OF HRA METHODS

This Appendix provides a short description of the HRA-Methods that are discussed and evaluated in Chapter 5 of this report. It contains:

A.1 THERP

A.1.1 Description of the error model

Boolean algebra is the main error model of the THERP method. In order to calculate the (initially unknown) failure probability of an operator task, the principle of action or task decomposition is used. It is assumed that this failure probability using the operation rules of Boolean algebra can be calculated from the single failure probabilities of the sub-tasks into which the task was previously decomposed. The 'tool' for task decomposition and subsequent error quantification is an event tree ("HRA event tree"), which models the logical and time-dependent correlations between the individual success and failure events. Figure A-1 shows such an event tree. It is an error model for a task (switching to the circulation mode) to be performed by the personnel in the course of a major loss-of-coolant accident in order to prevent failure of the pumps required for decay heat removal.

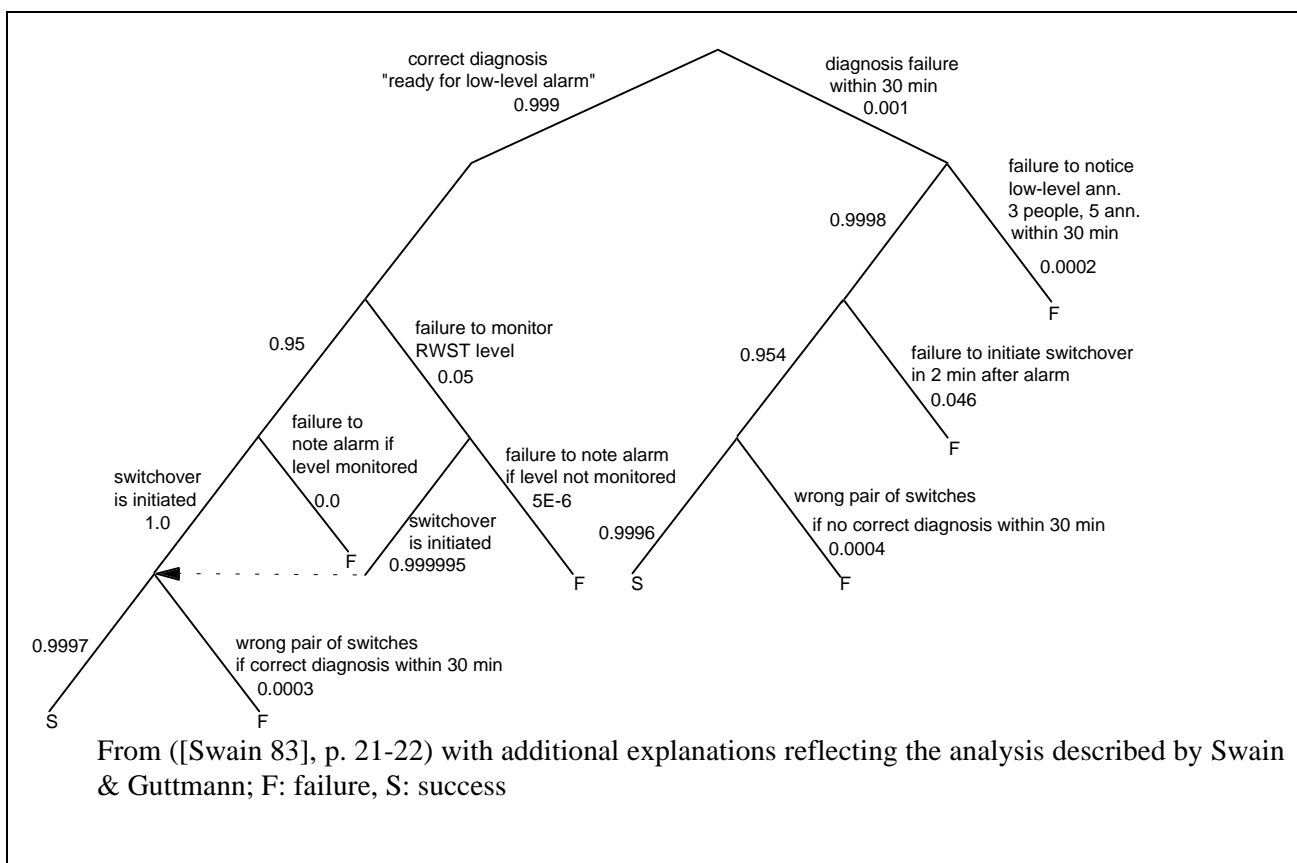


Figure A-1: THERP event tree modelling for the failure of the task of changing from feed to circulation mode as soon as the alarm warns of an excessively low water level in the refuelling water storage tank (RWST) during a large loss-of-coolant accident

The analysis example illustrates two principles to be applied in decomposing the task in the sense of THERP:

- the consistent consideration of dependencies between sub-tasks; cf. [Swain 83], Chapter 7;
- decomposition into meaningful sub-tasks, so-called perceptual units; cf. [Swain 83] (pages 11-5, 13-5, 15-14, etc.)

For the assessment of single error probabilities (HEPs) required for the quantification of an event tree for HRA, THERP uses different "submodels". The most important submodels are:

- time-reliability correlation to assess the probability for a diagnosis failure as a function of the time t available for diagnosis:

$$\text{HEP}(t) = \text{pr}(T > t),$$

where the time T required for diagnosis is modelled as a random variable;

- multiplicative model for adapting (modifying) a nominal error probability (NHEP) to conditions (e.g. stress) to be encountered in the plant or action situation and having a reliability-increasing or reliability-decreasing effect. The adaptation is carried out with the aid of influence factors, called "performance shaping factors" (PSFs) in THERP:

$$\text{HEP} = \text{NHEP} \cdot \text{PSF}_1 \cdot \text{PSF}_2 \dots,$$

where rarely more than two PSFs are assumed. Numerous PSFs (e.g. arrangement of switches) are already considered in selecting the NHEP.

A.1.2 Description of the database

The data contained in Swain's handbook (1983) mainly consist of:

- descriptions for approx. 100 NPP-specific errors with associated boundary conditions and
- numbers for point values and uncertainty bounds of the corresponding error probabilities.

Table A-1 gives a rough survey. In addition, Swain's handbook specifies a few factors (PSFs) for the situation-specific modification of error probabilities, e.g. due to stress or lack of operator experience. Table A-2 gives a rough survey concerning the sources of THERP data.

A.1.3 Description of performance shaping factors (PSFs)

An analysis carried out with THERP takes a number of performance shaping factors (PSFs) into consideration, which have a reliability-decreasing or -increasing effect. The PSFs specified in THERP can be roughly classified into four groups:

1. PSFs already included in selecting the nominal error probability (NHEP), e.g.: in a long written procedure the error of omission probability is 0.01 per step, in a short one 0.003 ([Swain 83], Table 20-7).

2. PSFs specified in the form of factors by which an NHEP is to be multiplied, e.g.: for medium-high stress the failure probability of a rule-based action increases by a factor of 2 ([Swain 83], Table 20-16).
3. PSFs specified as rules for modifying an HEP within their uncertainty limits, e.g.: use the upper bound of diagnosis failure probability if the diagnosis of the analysed accident is not regularly trained ([Swain 83], Table 12-5).

Table A-1: Important operator errors contained in the THERP database

<i>Operator error</i>	<i>NHEP(*UF) [Swain 83]</i>
1. incorrect accident diagnosis	0.00001(*/30) ... 1.0
2. misreading or failure to note information in observing system state displays	0.001(*/3) ... 0.1(*/5)
3. failure to respond to an audible alarm	0.0001(*/10) ... 0.25(*/10)
4. performance of a procedure without using the relevant written document	0.005(*/10) ... 0.3 (*/5)
5. incorrect use of a checklist	0.5 (*/5)
6. omission of a step in performing a procedure	0.001(*/3) ... 0.05(*/5)
7. error of selection in actuating a component (switch, hand valve) or in referring a system state display	0.0005(*/10) ... 0.05(*/5)
8. incorrect performance of a dynamic (knowledge-based) action under extremely high stress	0.25(*/5)

NHEP: nominal human error probability (median); UF: uncertainty factor (error factor);
For German translations, see [Reer 88]

Table A-2. Sources of Swain's handbook error probabilities

<ul style="list-style-type: none"> • Direct estimates of the handbook authors • Based on empirical data <ul style="list-style-type: none"> * nuclear power plant experience * simulator studies * industrial and military studies * "artificial" experiments * source not clear – Empirical probabilities directly adopted – Empirical probabilities modified with factors <ul style="list-style-type: none"> factors are estimates of the handbook authors factors are based on empirical data
--

4. PSFs which are mentioned, but for which no quantitative data are given ([Swain 83], Chapter 3).

The THERP user is generally urged to verify the transferability of the THERP HEPs to the respective analysis, using the PSFs discussed in Chapter 3 of the handbook, and to make modifications within the uncertainty bounds, if required ([Swain 83], page 3-10).

A.1.4. Description of operator decisions at the knowledge-based level

The THERP diagnosis model ([Swain 83], Chapter 12) is suitable for quantifying knowledge-based performance in accident diagnosis. The identification of an abnormal event to the effect of knowing with which component or system the problem caused by the event can be eliminated or reduced is defined as

diagnosis in THERP. The probability for a failure of the correct diagnosis is modelled in a time-dependent manner.

Actions after diagnosis are classified by Swain & Guttmann into the categories "step by step" (corresponding to rule- or skill-based) and "dynamic" (corresponding to knowledge-based). An increased error probability is recommended for knowledge-based actions, higher by at least a factor of 2.5 than for rule-based actions ([Swain 83], Table 20-16).

In addition, Swain's handbook also contains probabilities for errors whose origin is also influenced by knowledge-based performance elements, e.g. failure to notice that a measuring instrument displays a wrong reading (due to a technical defect), although the wrong reading can be detected by using further system state parameters ("cross check") ([Swain 83], Table 11-13). The THERP dependence model also allows the consideration of errors which greatly depend on knowledge-based performance, e.g. "... neither of the operators expects the other to make mistakes" ([Swain 83], p. 10-18).

A.1.5 Description of time reliability correlation (TRC)

In principle, a failure probability (HEP) of a manual action during accidents, calculated with THERP, is a time-dependent variable. The time dependence results from the THERP diagnosis model in which the estimated course (Table A-3) of the diagnosis failure probability $\Pr(T>t)$ is described as a function of the time t available for diagnosis. The total failure probability follows:

$$\text{HEP}(t) \approx \Pr(T>t) + \Pr(A)$$

where $\Pr(A)$ is the failure probability of post-diagnosis action(s). For the assessment of the complementary distribution function $\Pr(T>t)$, the diagnosis time modelled as random variable T , the THERP user may choose among three curves taking the education- and training-related knowledge of the control room personnel as a guideline (Table A-4): the better the knowledge about the accident to be analysed, the faster a correct diagnosis is to be expected.

The available diagnosis time t depends on the dynamic plant behaviour and on the time requirement t_A of the actions after diagnosis. This time requirement – in contrast to the diagnosis time requirement T – is not modelled as a random variable in THERP, but as a deterministic variable. Exercises or surveys already carried out are to be used for estimating t_A [Swain 83] (p. 6-11). ASEP [Swain 87] (Table 8-1) specifies estimates which can be used for determining t_A if no empirical time data are available. The following relation then often results for the available diagnosis time (t)

$$t = t_{\text{total}} - t_A,$$

where t_{total} is the total permissible time from the start of an accident to success of the required manual action.

For certain cases of analysis, the dynamic plant behaviour is the main criterion for estimating t . An example is shown in Figure A-1 analysing a change to the circulation mode. This action must be taken 2 min, at the latest, after the alarm warns of an excessively low refuelling water storage tank level. The alarm is to be expected 30 min after the occurrence of an accident. These 30 min are assumed in [Swain 83] (p. 21-12) as the available diagnosis time, which means that 2 min are assumed to be sufficient for the required switching action.

The probabilities in Table A-3 are applicable if only *one* abnormal event occurs. If a second or third event occurs, THERP assumes that the time required for diagnosis is extended by 10 min for each additional event; this estimate is based on simulator studies by [Woods 82]. For two events, for example, a median (MDN) of 0.1 (instead of 0.01 for one event) would result for $t = 20$ min, for three events the median would be $MDN = 1.0$. This 10-minute rule is further illustrated in ASEP [Swain 87] (Table 8-1).

Table A-3. Time-dependent parameters of diagnosis failure probability

t	1 min	10 min	20 min	30 min	60 min	1500 min
MDN(t)	1.0	1.0E-1	1.0E-2	1.0E-3	1.0E-4	1.0E-5
LB(t)	1.0	1.0E-2	1.0E-3	1.0E-4	3.3E-5	3.3E-6
UB(t)	1.0	1.0	1.0E-1	1.0E-2	3.0E-3	3.0E-4

MDN = median; LB = 5 % lower bound; UB = 95 % upper bound; t = available diagnosis time. The parameters between two t-values are calculated by a loglinear interpolation. After [Swain 83] (Table 20-3).

Table A-4. Guidelines for assessing the time dependence of diagnosis failure probability.

<ol style="list-style-type: none"> 1. Use upper bound (UB) if <ol style="list-style-type: none"> a. the situation is not covered in training or b. the situation is covered but not practised except in initial training of operators for becoming licensed, or c. the talk-through and interviews show that not all the operators know the pattern of stimuli associated with the situation. 2. Use lower bound (LB) if <ol style="list-style-type: none"> a. the situation is a well-recognised classic incident and b. the operators have practised this situation in simulator requalification exercises, and c. the talk-through and interviews indicate that all the operators have a good recognition of the relevant stimulus patterns and know which actions (procedures) to follow. 3. Use median (MDN) if <ol style="list-style-type: none"> a. the only practice for coping with the situation is simulator requalification exercises and all operators have had this experience, or b. rules 1 and 2 above do not apply.

From [Swain 83] (p. 12-23). The guidelines relate to the time dependencies outlined in Table A-3. Further guidelines are contained in ASEP ([Swain 87], Table 8-1), e.g. the conditions under which a reduced diagnosis probability (lower bound) may be assumed in the case of symptom-oriented procedures.

Moreover, it should be noted that, in addition to TRC, THERP gives guidelines to quantify some other aspects of time, especially in periodically scanning or resuming attention to an alarm; see [Swain 83] (Tables 20-24 and 20-25).

A.1.6 Description of dependencies between errors

The model recommended in THERP for the quantification of dependencies between events relating to human action corresponds to the beta factor model known from technical reliability analysis. The following applies to the conditional occurrence of an event B:

$$\text{pr}(B|A) = \left[\frac{1 + (n-1)\text{pr}(B)}{n} \right] = \beta + (1-\beta)\text{pr}(B)$$

where:

$\text{pr}(B|A)$ = probability for B on condition that A has occurred;

$\text{pr}(B)$ = probability for the occurrence of B independently of A, also denoted as basic human error probability (BHEP);

$\beta = 1/n$ = level of dependence (beta factor), corresponds to the probability that the cause for the occurrence of A will also lead to the occurrence of B;

$n = 15$ ($\beta = 0.05$) for low, $n = 7$ ($\beta \approx 0.15$) for moderate, $n = 2$ ($\beta = 0.5$) for high and $n = 1$ ($\beta = 1.0$) for complete dependence between A and B.

All levels of dependence recommended in the model are not based on empirical data, but on a subjective discretisation of the value range from $\beta = 0$ (no dependence) to $\beta = 1.0$ (complete dependence). Swain & Guttmann give a number of guidelines for selecting the level of dependence. The most important criteria are summarised in Table A-5.

The quantification of personnel redundancy in an accident sequence is one of the most frequent applications of the dependence model. THERP recommends here the assumptions summarised in Table A-6.

The discretisation of the steady range from $\beta = 0$ to $\beta = 1$ into five discrete points in THERP facilitates application. The resultant inexactness is to be regarded as insignificant in view of the data problem and the PRA-typical uncertainties. Particularly helpful for the THERP user are the guidelines which THERP gives to assess the level of dependence (Table A-5).

Table A-5: THERP guidelines for assessing the level of dependence between two tasks or acting persons (after [Swain 83], Chapters 10 and 18).

ZD (zero dependence)	Rare, unusual, time (> 1 min) and spatial differences between 2 tasks. Example: verification of 2 instruments at different locations in the control room during normal operation.
LD (low dependence)	If there is any doubt as to an assessment of independence (ZD); slight time (approx. 1 s to 60 s) and spatial differences between 2 tasks. Examples: between 2 operators newly acquainted with each other; between shift supervisor and new operators; between shift supervisor and other operators, if the system status must be assessed after a transient.
MD (moderate dependence)	Obvious relationship between 2 tasks; slight time (approx. 1 s to 60 s) and spatial differences. Example: between shift supervisor and other operators for tasks interacting with each other.
HD (high dependence)	Substantial relationship between 2 tasks; very slight time (approx. 1 s) and spatial differences. Examples: operator A significantly differs from operator B with respect to prestige and authority, so that the behaviour of A has a great influence on the behaviour of B; between shift supervisor and reactor operator.
CD (complete dependence)	Rare, 2 tasks are practically performed simultaneously in space and time. Example: operation of paired switches.

Table A-6. THERP assumptions on personnel available to cope with an accident and existing levels of dependence (after [Swain 83], Table 20-4).

Time after first response to an abnormal event (accident)	(a) Personnel available to cope with the accident	(b) Dependence levels
(1) 0 to 1 min	one reactor operator (RO)	
(2) \approx 1 min	one reactor operator (RO) shift supervisor or deputy	HD with RO
(3) \approx 5 min	one reactor operator (RO) experienced senior operator shift supervisor one (several) assistant operator(s)	HD with RO LD to MD with others
(4) \approx 15 min	one reactor operator (RO) experienced senior operator shift supervisor shift technical advisor one (several) assistant operator(s)	HD with RO LD to MD with others LD to MD with others for diagnosis and major accidents HD to CD for detailed operations

A.1.7 Description of error correction (recovery)

The THERP method permits the quantification of a number of possibilities for correcting an operator error ([Swain 83], Figure 20-1):

1. personnel redundancy: the acting operator is checked by another operator;
2. subsequent procedure step or task;
3. annunciating display (audible alarm);
4. periodical scanning inside control room (CR) or walk-around inspection outside CR.

The THERP model assumes that safety-relevant deviations from the nominal status can be detected by CR scanning or walk-around inspections during normal operation. To this end, a number of situation-specific error probabilities are listed which, among other aspects, depend on the respective rounds strategy [Swain 83] (pages 11-25, 11-29, 19-15, 19-22). For the quantification of personnel redundancy in normal operation THERP provides a number of situation-specific control error probabilities ranging between 0.001 and 0.5 [Swain 83] (Table 20-22). In the event of an accident, THERP recommends that personnel redundancy should be quantified through the dependence model [Swain 83] (Table 20-18). This model should also be used for the quantification of recoveries by subsequent tasks; see [Swain 83] (page 10-7).

According to THERP, the response to an audible alarm is a recovery factor for the failure of diagnosis in an accident. Swain & Guttmann recommend a so-called alarm-response model in which the error probability depends on the total number of alarms and the place where the alarm considered as recovery factor ranks [Swain 83] (Table 20-23). Several alarms may have to be combined to so-called perceptual units. Moreover, the failure to notice an alarm is to be quantified considering the personnel redundancy existing in the control room. Figure A-1 shows an example of the quantification of the recovery potential

of an alarm, as recommended in THERP. The probability (0.0002) that the diagnosis failure (0.001) is not recovered results from the product of the probabilities

- 0.003 for an operator's failure to notice a particular alarm out of a total of 5 alarms,
- 0.5 for failure to check by a second operator with high dependence on the preceding error and
- 0.15 for failure to check by a third operator with moderate dependence on the preceding errors.

The diagnosis failure probability is thus reduced from 0.001 to $0.001 \cdot 0.0002 = 2E-7$ due to the alarm.

A.1.8 Description of uncertainties

Probabilities for the evaluation of human reliability are modelled by THERP as uncertain variables. For this reason, an uncertainty interval (lower bound (LB); upper bound (UB)) is also specified for a probability in addition to a point value (HEP). The interval should contain the true value of the probability with a high certainty (90 %). The following uncertainties are to be covered by these intervals according to Swain & Guttman [Swain 83] (p. 7-9):

1. random fluctuations in human reliability;
2. imperfect knowledge of the THERP user.

These two main groups are specified in more detail in Table A-7. The uncertainty is mathematically modelled by a lognormal distribution. An estimated point value of an error probability is defined as the median (MDN) of this distribution. The uncertainty is expressed by a scattering factor (K) and - according to the lognormal distribution - quantified by a multiplicative model:

$MDN = LB \cdot K$, $UB = MDN \cdot K$, with the restriction that the value 1.0 must not be exceeded. Details on the mathematical operations can be found in Appendix A of Swain's handbook [Swain 83].

Table A-7. Variability causes to be covered by the uncertainty factors in Swain's handbook. From [Swain 83] (pages 7-9 to 7-10)

- | |
|--|
| <ol style="list-style-type: none"> 1. Dearth of data on human performance in NPPs that is in a form suitable for HRA. 2. Inexactness of models of human performance that purport to describe how people act in various situations and conditions. 3. Inadequate identification of all relevant performance shaping factors, of their interactions and effects. 4. Limitation in the skill and knowledge of the human reliability analyst. 5. Variability in performance within an individual and among the performances of different individuals. |
|--|

A.2 Brief Description of EDFs PHRA

EdF distinguishes in the same way as THERP between diagnosis and execution of post-diagnosis actions. In contrast to THERP, EdF considers the possibility that the accident can be successfully controlled even with an incorrect diagnosis; however, the corresponding conditional error probability is rated very high. THERP quantifies this error probability with 1.0: if no correct diagnosis, then failure of accident control by the personnel. In order to assess the diagnosis failure probability as a function of available time, EdF distinguishes between four time curves (Figure A-2):

- two pessimistic time curves based on the THERP diagnosis model (curves 2 and 3),
- two less pessimistic time curves based on French simulator data (curves 1 and 1'),

where the complexity of the accident situation is to be taken as the selection criterion.

According to the EdF model concept, there is a so-called "residual probability" in the diagnosis; this is the probability that the correct diagnosis will never be made by the crew itself without any help (from the safety engineer or the technical support center). This probability is 0.03 (in the case of extremely difficult accident diagnosis, curve 3 in Figure A-2) or 0.005 (in other cases, curves 1, 1' and 2), see [EPS 1300] (page 112).

The EdF method furthermore provides for the possibility of quantifying some kinds of misdiagnoses leading to unrequired actions (unwarranted shutdown of safety injection after LOCA; unwarranted isolation of steam dump to the atmosphere after steam generator tube rupture) [Mosneron 94]. For these specific cases, the execution model (see below) can be used (though it is generally used for "pure" execution, not including diagnosis) [Mosneron 95].

Every important (critical) action after diagnosis is to be quantified with the so-called "execution model" [Mosneron 90] (Section 2.3.3.3):

$$P_E = P_B \cdot K_F \cdot P_{NR}$$

where:

P_E = probability that execution of action fails;

P_B = basic probability for failure;

K_F = modification factor with respect to the action context;

P_{NR} = recovery failure probability.

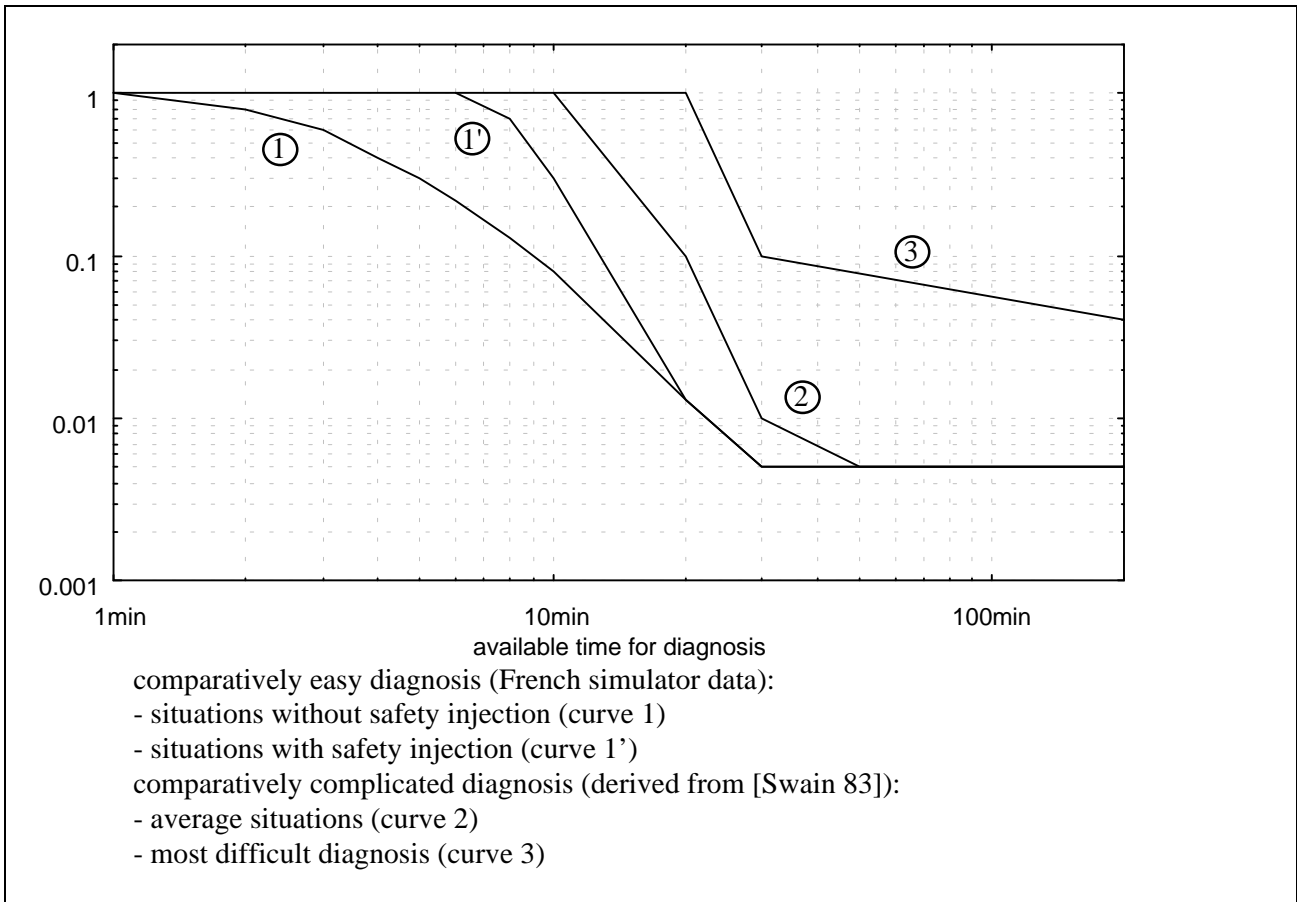


Figure A-2: Time-dependent curves of diagnosis failure probability recommended in EdF's PHRA [EPS 900] (Table 5-20).

For the model parameters (P_B , K_F , P_{NR}) situation-specific numbers are given. These numbers are based on a combination of simulator data and expert judgement.

If a reliable estimate on the average time requirement is available for the entire action (including diagnosis), then the time-dependent error probability is to be determined according to standardised (i.e. normalised) curves outlined in Table A-8. In this case, the following additional contributions (see above) to failure of the manual action are to be quantified:

- residual probability that the diagnosis will never be made;
- probability that the execution after diagnosis fails.

$t/T_{0.5}$	curve 1	≈ 0.3	≈ 1	≈ 2.5	≈ 3.2	≈ 4.8
	curve 2			≈ 3.1	≈ 4	≈ 7.6
$pr(T > t a)$		1.0	0.5	0.1	0.05	0.01

Table A-8. Some typical values of the probability $pr(T > t|a)$ that the action is not performed within t on condition that the action is performed at all (derived from [Mosneron 90] (Figure A-2). In selecting the curve, the complexity of the situation and the availability of experimental results must be taken into consideration [EPS 1300] (see 6.4.1).

These two approaches - (1) TRC via diagnosis curve, (2) TRC via standardised curve - may be illustrated by the following example (taken from [EPS 1300], pages 111-113):

Accident

Shutdown state, break in primary system, safety injection system (SIS) fails to start up automatically.

Action

Operators bring SIS into operation.

Analysis (1) Incorporation of TRC via diagnosis curve

$$p = p_D p_{E|D} + (1 - p_D) p_{E|d} = 0.005 \cdot 1 + (1 - 0.005) \cdot 0.006 \cong \underline{0.011}$$

where:

p = total failure probability;

$p_D = 0.005$ = probability of diagnosis failure; curve 2, $t = 50$ min available;

$p_{E|D} = 1$ = conservative assessment of the probability that execution of action fails, given diagnosis failure;

$p_{E|d} = p_B K_F p_{NR} = 0.06 \cdot 1 \cdot 0.1 \cong 0.006$ = probability that execution of action fails, given correct diagnosis;

$p_B = 0.06$ = basic execution failure probability; to choose between 0.02 and 0.06, according to the type of error;

$K_F = 1$ = factor for contextual correction; to choose between 1, 1/3, and 3, depending upon the context in question;

$p_{NR} = 0.1$ = probability of non-recovery; to choose between 0.03, 0.1, 0.3, and 0.6 depending on the recovery factors (explicit human redundancy and/or alarm signal) and available recovery time (<30min or >30min).

Analysis (2) Incorporation of TRC via standardised (normalised) curve

$$p = p_{RA} + (1 - p_{RA}) p_{A(t)|a} = 0.007 + (1 - 0.007) \cdot 0.002 \cong \underline{0.009},$$

where:

$p_{RA} = p_{RD} p_{RE|D} + (1 - p_{RD}) p_{RE|d} = 0.007$ = residual probability that no action will be implemented;

$p_{RD} = 0.005$ = residual probability that accident diagnosis will never take place; to choose between 0.03 (difficult diagnosis) and 0.005 (other cases);

$p_{RE|D} = 1$ = conservative assessment of the residual probability that the action will never be implemented, given that the diagnosis has failed;

$p_{RE|d} = p_B K_F p_{NR} = 0.06 \cdot 1 \cdot 0.03 \cong 0.0018$ = residual probability that the action will never be implemented, given that the diagnosis has been successful; where $p_{NR} = 0.03$ (instead of 0.1) since it is a residual probability;

$p_{A(t)|a} = \text{pr}(T > t|a) = 0.002$ = probability of failure to implement action within time limit t , given that action will eventually be implemented; from standardised curve 1 at $t/T_{0.5} = (50 \text{ min}) / (7 \text{ min}) \cong 7$.

The intervention by an outside expert, the so-called "safety engineer", is quantified by EdF as an additional opportunity for recovery. The safety engineer is not permanently present in the control room, but can (and should) be called in the event of an accident. The failure probability of an error recovery by

the safety engineer is modelled dependent on time and situation [EPS 1300] (Chapter 5.3.3.5). The required data are based on simulator tests and on the dependence model from [Swain 83]. Finally, the intervention of the technical support centre is very roughly modelled considering that no important error occurs later than four hours after the beginning of the accident ([EPS 1300], page 107).

EdF's model of recovery by the safety engineer may be illustrated in the following example (taken from [EPS 1300], page 115).

Accident

Shutdown state, break in primary system, safety injection system (SIS) fails to start up automatically, operators (without safety engineer) fail to bring SIS into operation.

Action

Safety engineer recovers operators' failure.

Analysis

$$p = A(t) + [1 - A(t)] p_{E|a} = 0.15 + (1 - 0.15) \cdot 0.05 \cong 0.2,$$

where:

p = probability that recovery by the safety engineer fails;

$A(t) = 0.15$ = probability of absence of the safety engineer at a given time $t = 50$ min (time required for corrective action is assessed as negligible) beyond which it is too late to implement corrective action; $A(t)$ is obtained from the simulator tests and plant-specific survey;

$p_{E|a} = 0.05$ = probability that recovery fails when the safety engineer is present; to choose between 0.05, 0.1, and 0.3, depending on the type of written instructions for the corrective action and the time available to implement the corrective action; the numbers are derived from simulator tests and the dependence model in [Swain 83].

Furthermore, EdF's PHRA includes a method for assessing the uncertainty of an estimated HEP point value, see [EPS 1300], Section 5.3.1.6. Finally, the reader should take notice that EdF's methodology has been updated in 1995 to account for symptom-based procedures and the new organisation of the crews [Mosneron 96]. The main changes are:

- The standardised curves (see Table A-8 in this report) will be no longer used.
- The guidelines for assessing the probability (p_{NR}) of non-recovery (by the crew) have been adapted to the structure of symptom-based procedures (with new recovery factors).
- The HEPs ($A(t)$, $p_{E|a}$) of the safety engineer model have been reduced to account for improvements in the organisation.

A.3 Brief Description of HCR

The HCR model calculates the failure probability, HEP, for a task with the aid of a three-parameter Weibull distribution $Q(t)$, Figure A-3. The following two estimates are needed to determine the three parameters uniquely determining the course of $Q(t)$:

1. Cognitive demand on the personnel by the task broken down into skill-based, rule-based and knowledge-based requirements after [Rasmussen 79].

2. Point estimate $T_{0.5}$ (median) of the average time required for performing the task. If no empirical time data are available for the determination of $T_{0.5}$, then $T_{0.5}$ is to be estimated by expert judgement (supported by plant visits and operator interviews). A value determined for $T_{0.5}$ has still to be modified upwards (at most by 300 %) or downwards (at most by 40 %) using three PSFs (see Figure A-3) in order to include the boundary conditions to be expected in the action situation.

The failure probability will then result by substituting the time t available for performing the task in the complementary distribution function, $P(t) = 1 - Q(t)$, of the Weibull distribution determined before.

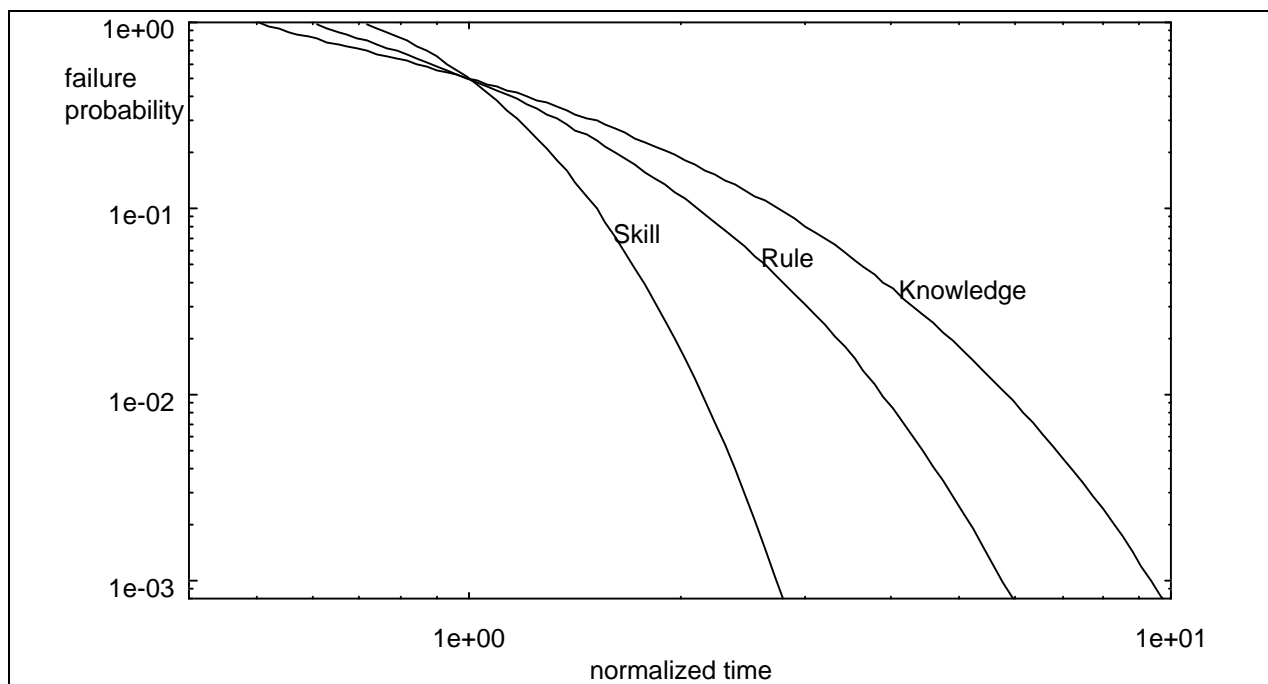


Figure A-3: Dependence of the error probability p on the quotient of available time (t) and required average time ($T_{0.5}$), as quantified in the HRC model [Hannaman 84] for the performance of a task.

The error is designated "non-response" in the HCR model and defined as non-completion of a predefined task. The average time required, $T_{0.5}$ (median time), results from a nominal estimate ($T_{0.5}/\text{nominal}$) still to be modified by three PSFs:

$$T_{0.5} = T_{0.5}/\text{nominal} (1 + K_1) (1 + K_2) (1 + K_3)$$

where:

K_1 = coefficient for operator experience, $K_1=0.22$ (expert, well-trained), $K_1=0$ (average knowledge training), or $K_1=0.44$ (novice, minimum training);

K_2 = coefficient for stress level, $K_2=0.44$ (situation of grave emergency), $K_2=0.28$ (situation of potential emergency), $K_2=0$ (active, no emergency), or $K_2=0.28$ (low activity, low vigilance);

K_3 = coefficient for quality of operator/plant interface, $K_3=-0.22$ (excellent), $K_3=0$ (good),

$K_3=0.44$ (fair), $K_3=0.78$ (poor), or $K_3=0.92$ (extremely poor).

The time-dependent error probability is:

$$p(t) = \exp\{-(t/T_{0.5} - a)/c\}^b$$

where:

- (a, b, c) = behaviour-type-specific coefficients,
 (0.7, 1.2, 0.407) for skill-based behaviour,
 (0.6, 0.9, 0.601) for rule-based behaviour, or
 (0.5, 0.8, 0.791) for knowledge-based behaviour.

The HCR model is to be seen in connection with the SHARP method. This method is essentially based on action decomposition, as is THERP. In contrast to THERP, action decomposition in SHARP follows a slightly coarser pattern similar to the OAT method ("Operator Action Tree") [Hall 82], and greater emphasis is placed on modelling time-dependent error probabilities. Figure A-4 illustrates that the HCR model is primarily conceived for the quantification of cognitive processes in accident diagnosis:

- In particular, the model considers different levels of thinking or cognitive processing associated with the way the crew supervises the course of an accident sequence« [Hannaman 85] (page 343) .

In concrete terms, the HCR model is to furnish a probability for the error designated A3 in Figure A-4: no detection or slow cognitive processing relative to the requirements of the event. For the other errors associated with the diagnosis,

- (A2) misdiagnosis: faulty detection/cognitive processing,
- (B4) selection of a non-viable option despite correct identification of the accident-initiating event,
- (B6) selection of a non-viable option in the case of misdiagnosis and
- (C) recovery of a preceding diagnosis or decision-based error fails,

the HCR model refers to other methods and data sources. Errors in connection with the performance of an action (after correct diagnosis and decision) are not dealt with in the model in Figure A-4.

As stated elsewhere, such errors belong to the "other" aspects (not covered by HCR) of an analysis:

»The non-response probability« (from the HCR model) »can be incorporated into the logic representation defined in step 4« (representation) »of SHARP (e.g. EOATS, HRA trees) to include other aspects of human reliability such as turning the wrong switch, given that the intent was to turn the correct switch « [Hannaman 85] (page 348).

An error probability calculated with the HCR model is thus a contribution to the failure probability of diagnosis and decision-making after the occurrence of an accident. Guidelines for assessing the uncertainty bounds are given in [Moieni 86].

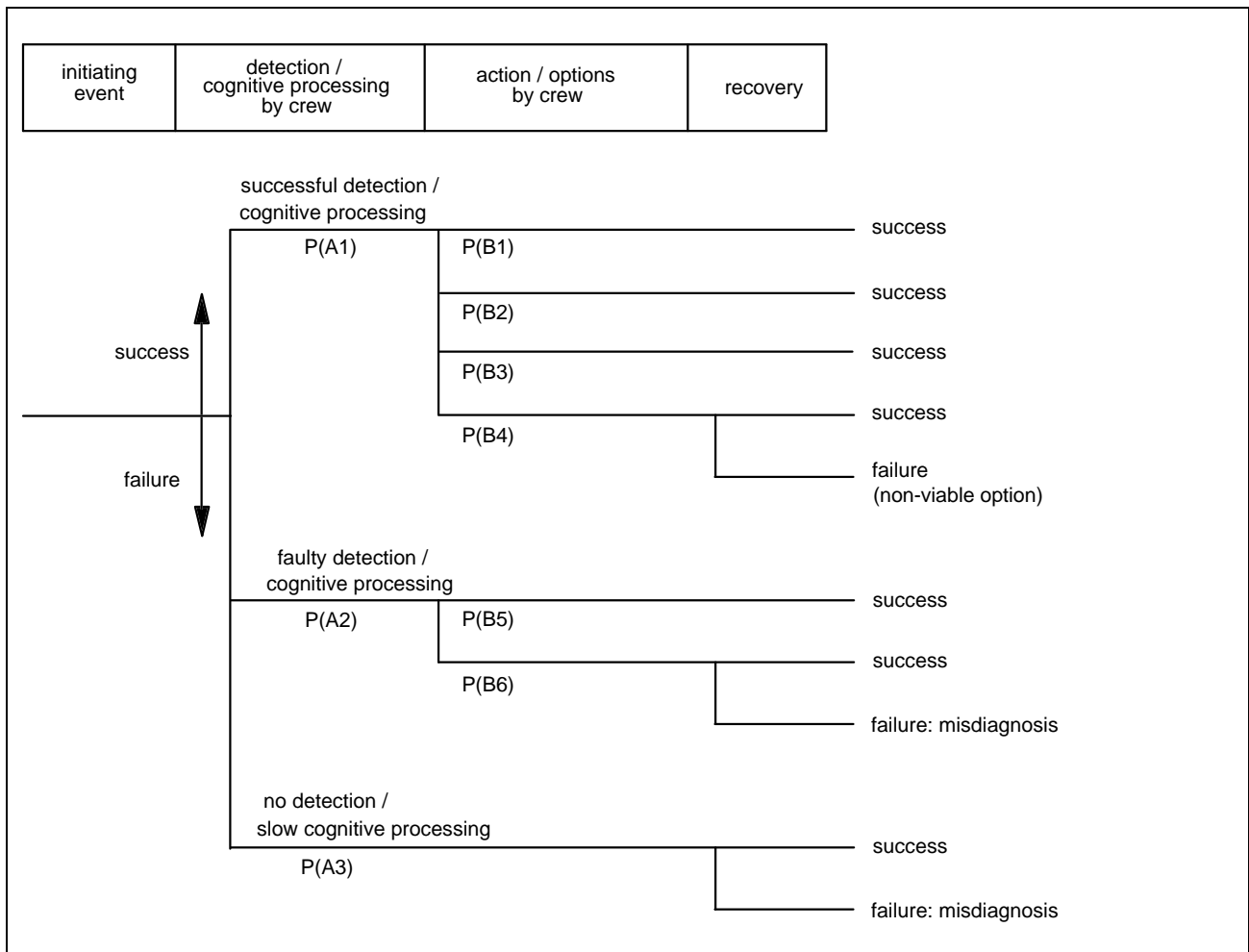


Figure A-4: Expanded operator action tree [Hannaman 85] for the incorporation of the HCR model. The HCR model itself (Figure A-3) serves to estimate P(A3).

A.4 Brief Description of HCR/ORE

The error model for actions with predefined procedure is outlined in Figure A-5. A distinction is made between diagnosis and subsequent actions. THERP [Swain 83] is referred to for the quantification of actions after diagnosis ([Moieni 94], p. 41).

Two main types of error are distinguished as contributions to diagnosis failure:

1. failure to formulate a correct response
2. failure to initiate a correct response in a timely manner (too late) given that a correct response was found.

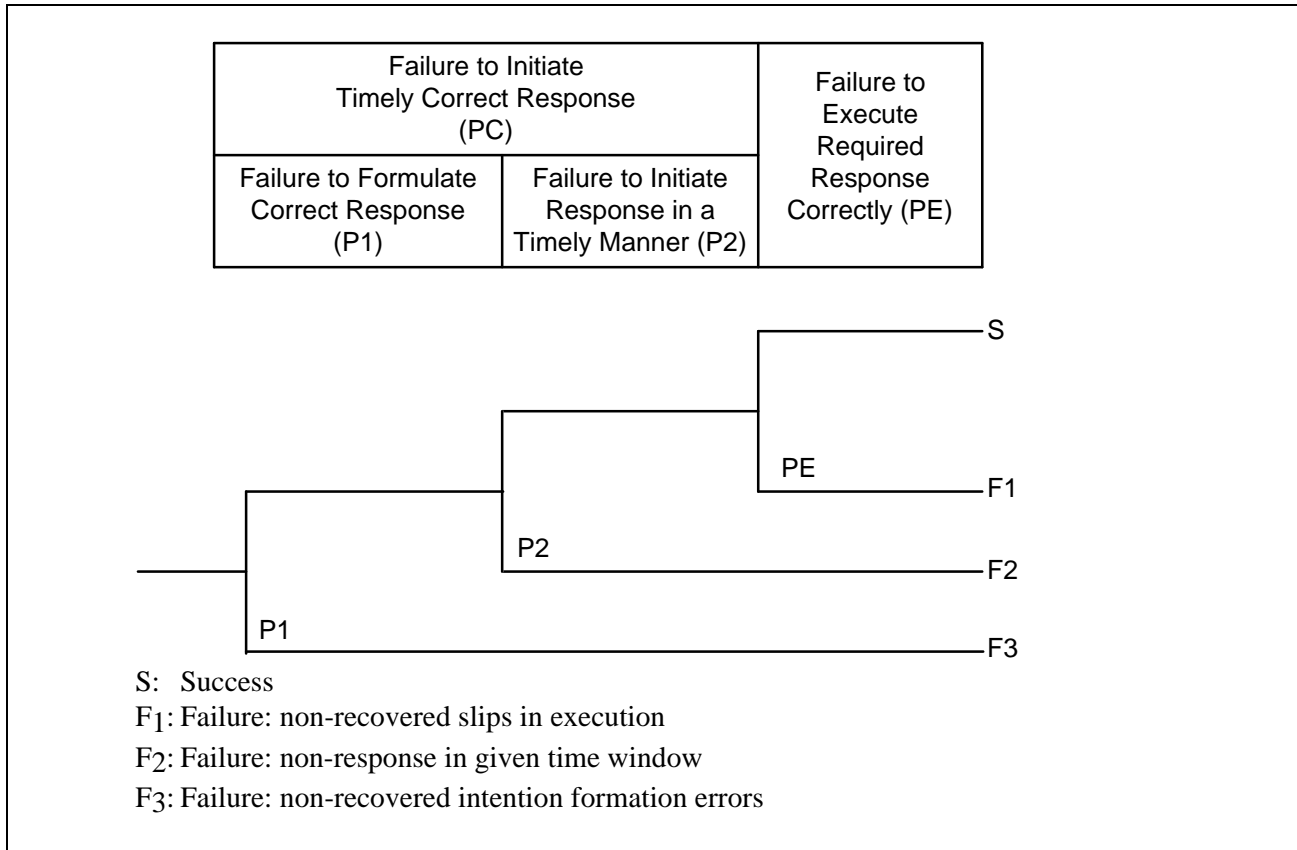


Figure A-5: Generalised event tree for modelling procedure-driven operator actions in an accident ([Moieni 94], Fig. 4)

Eight error mechanisms are identified ([Moieni 94], p. 40) for the first error type (failure to formulate correct response):

- a) problems due to available information;
- b) failure of attention;
- c) misreading/miscommunication;
- d) misleading information;
- e) skipping a procedure step;
- f) misinterpretation of instruction in procedure;
- g) misinterpretation of decision logic in procedure;
- h) deliberate deviation from the procedure (violation).

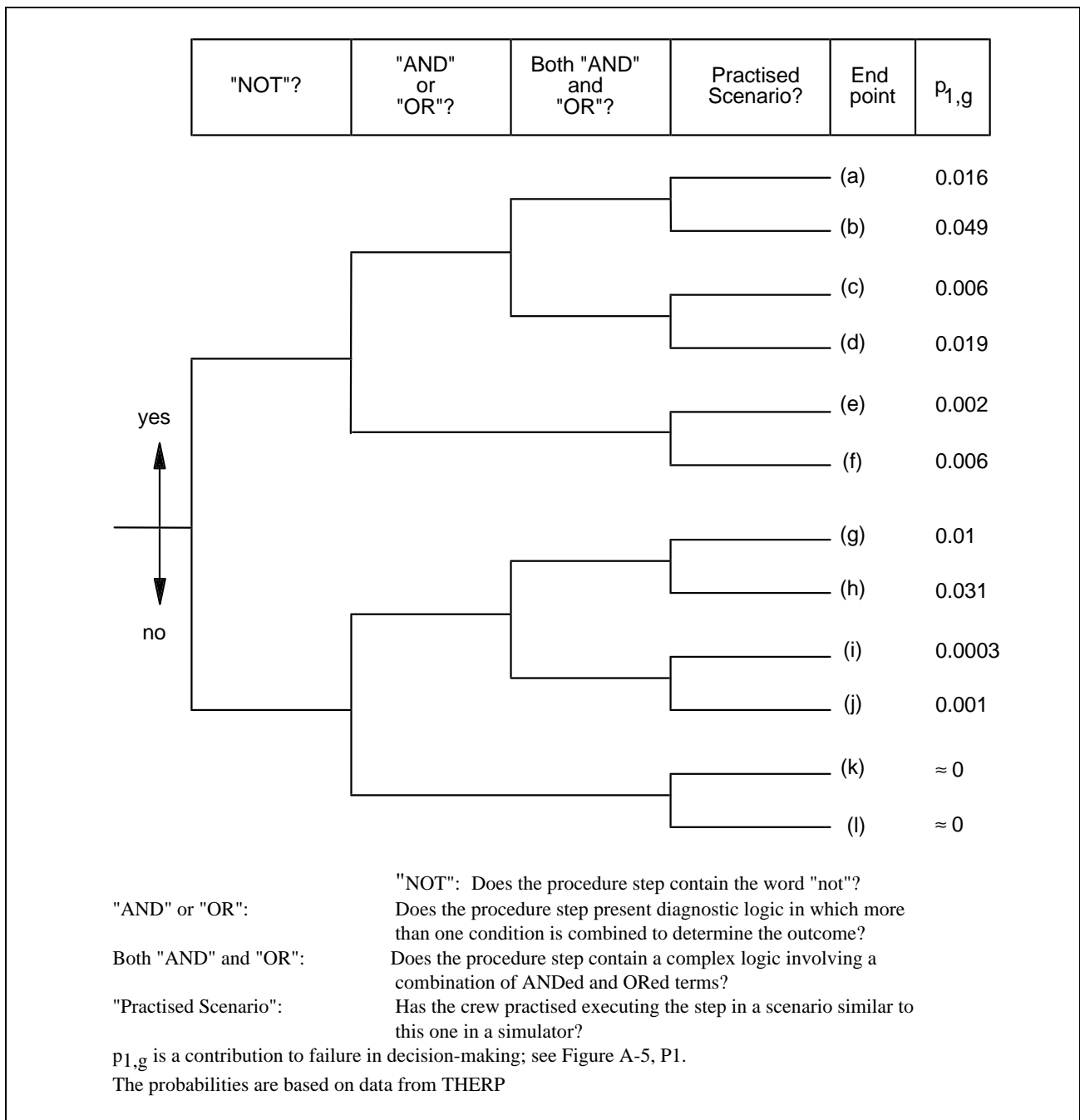


Figure A-6: Decision tree for determining the probability $p_{1,g}$ that a diagnostic logic in a procedure is misinterpreted ([Moieni 94], Figure 7)

For each mechanism, an error probability is to be determined with the aid of structured expert estimates. The decision tree method outlined in Figure A-6 is recommended for this purpose. Uncertainty is not explicitly addressed; since the HEPs obtained from the decision trees are based on THERP, they can be treated as median values of lognormal distributions with error factors ranging from three to ten [Moieni 94] (page 40). Furthermore, recovery mechanisms are to be considered for each error mechanism, see [Moieni 94] (page 41) for details.

The second error type (late diagnosis) in Figure A-5 is to be quantified with the aid of time-reliability correlations from the ORE program ([Moieni 94], p. 38). The ORE data were processed according to the same methodological principle as the data of the HCR model – as complementary distribution functions of normalised times, normalising being effected by division by the respective average time requirement $T_{0.5}$ (median). The ORE diagnosis times were approximated with the aid of two-parameter lognormal distributions:

$$p(t) = \text{pr}(T > t) = 1 - \Phi\left(\frac{\ln(t / T_{0.5})}{\sigma}\right)$$

with:

T = diagnosis time (random variable);

t = available diagnosis time

σ = standard deviation from $(\ln T)$;

$\Phi(..)$ = distribution function of the standard normal distribution

This involves a total of 1,100 diagnosis times recorded under the ORE program ([Moieni 94], p. 39). In attempting to group these data with respect to typical σ -values, the HCR grouping according to cognitive levels (skill/rule/knowledge) proved unsuitable. Instead, six typical courses were identified using other criteria. First of all, two reactor-type-specific main types were identified, one for PWR accidents and one for BWR accidents. Within each reactor type group three time profiles are distinguished depending on the accident dynamics ([Moieni 94], page 33).

Table A-9 shows the six different types of accident diagnosis. The standard deviations (σ -values) of the logarithmized diagnosis times were determined here (in compiling Table A-9) by the following equation

$$\sigma = \frac{\ln(K_{0.95})}{1.645}$$

where the quotient, $K_{0.95} = T_{0.95}/T_{0.5}$, of 95% percentile ($T_{0.95}$) and 50% percentile ($T_{0.5}$) was read from the time profiles printed in [Moieni 94] (p. 38). Due to the readings performed here, the σ -values listed in Table A-9 are not particularly precise. The precise σ -values are the property of EPRI.

For the quantification of actions without predefined procedure (recovery actions), i.e. repair actions in the widest sense, the HCR/ORE method roughly distinguishes between

- identification of the recovery action (by the crew) and its
- execution,

where it may be necessary to subdivide the execution even more precisely. If there are empirical time data available for these sub-tasks, then the recovery action is to be modelled as a stochastic process, and the result obtained is a time-dependent error probability ([Moieni 94], page 45). Essentially, this corresponds to the approach proposed in [Reer 93] (Section 2.4.1) and [Reer 94b].

If no empirical time data are available, structured estimates are to be used for quantification and the user is offered two decision trees as a tool. The first tree serves to quantify the error probability p_I in identifying (i.e. diagnosis by the crew) the recovery action, it comprises 6 PSFs; the second tree comprises 4 PSFs and serves to assess the probability p_E for failure to execute the recovery action ([Moieni 94], pages 47-49):

(I) PSF assessments needed for a simplified estimation of p_I :

1. Is the recovery action identified in emergency operating or other procedures?
2. Are there direct indications in the main control room?
3. Is the time available for diagnosis long, intermediate or short?
4. Are the crew familiar with or trained on the recovery process?
5. If reasoning is required, is it deductive or inductive?
6. Is detection by auxiliary operators possible?

(E) PSF assessments needed for a simplified estimation of p_E :

1. Is the time for execution long, intermediate or short?
2. Are the crews trained (or practised) in the execution of the recovery action and/or is there a procedure?
3. Is the recovery action simple or complex?
4. Are the environmental conditions around the equipment hostile?

Finally, the non-recovery probability is

$$PNR = P_I + P_E - P_I P_E$$

Required response dictated by the dynamics of the accident	Reactor type	σ
(1) Response following a change in the plant state that is indicated by an alarm or value of a monitored parameter (e.g. response to a spurious pressurizer spray operation in a PWR).	PWR BWR	≈ 0.55 ≈ 0.73
(2) Response following an event that gives rise to a primary cue (as in case 1 above) that has to be achieved when a parameter is exceeded or can be seen not to be maintainable below a certain value (e.g. initiate residual heat removal when the suppression pool (SP) temperature exceeds 35 °C in a BWR). This human interaction involves a waiting period after the primary cue in order to reach a determined plant state.	PWR BWR	≈ 0.36 ≈ 0.61
(3) Response following an event that gives rise to a primary cue (as in case 1 above) that has to be achieved before some plant parameter reaches a critical value (e.g. initiate standby liquid control system before SP temperature reaches 43 °C in a BWR). This critical value can be regarded as a soft prompt, or secondary cue.	PWR BWR	≈ 0.76 ≈ 0.79

Table A-9. Standard deviation (σ) of the logarithmic diagnosis time ($\ln(t)$) as a function of response and reactor type. After [Moieni 94] (Figures 3 and 6). Except for PWR response type 3, the curves are plotted together with their uncertainty bounds.

A.5 SLIM

The SLIM method begins with the identification of relevant performance shaping factors which are of significance for the failure or success of an operator task to be analysed; for example, for the task of diagnosing the system state in the event of a transient and intervening if a pump of the emergency feedwater system fails ([Embrey 84], Vol. I, p. 5):

- quality of information available to the operators in the control room (PSF 1);
- quality of procedures (PSF 2);
- time available for diagnosis and performance of an action (PSF 3);
- training level of the operators (PSF 4).

The weighted sum of these PSFs then gives a reliability index SLI ("success likelihood index") defined in the interval (0;1):

$$SLI = \sum_{i=1}^n (w_i x_i),$$

with

$$\sum_{i=1}^n w_i = 1$$

where x_i is the degree of the presence (with respect to a reliability-increasing effect) of PSF i in the action situation investigated, w_i is the relative significance (weighting factor) of PSF i and n is the number of relevant PSFs.

The values for w_i and x_i also defined in the interval (0;1) are to be determined by expert judgement. In this connection, the method description repeatedly refers to the estimation techniques known from multi-attribute utility theory (MAUT); e.g. [Embrey 84] (Vol. I, page 5): »simple multi-attribute rating technique (SMART)« [Edwards 77].

There are also applications of SLIM in which a failure index (FLI) instead of a reliability index (SLI) is assessed; e.g. in the Mühleberg PSA. In [Reer 93] (Section 5.1.3, English version in [Reer 94c]) a method for the quantification of decision-based errors was developed on the basis of the FLI variant of SLIM.

The following loglinear model equation is recommended in [Embrey 84] (Vol. I, page 7) for the SLI-dependent calculation of the success probability q and the failure probability p of a task examined:

$$\log(q) = \log(1 - p) = a \cdot SLI + b$$

where at least two reference values are required for q and p (with associated SLI estimates) to determine the model parameters a and b . If more than two reference values are available, then a and b are to be determined by regression.

In other passages of the method description, the logarithmized error probability p is also modelled as a linear function of the reliability index SLI; cf. [Embrey 84] (Vol. I, page 14) or [Embrey 83] (page 15):

$$\log(p) = a \cdot \text{SLI} + b$$

Guidelines for uncertainty bounds estimation are given in [Embrey 84] (Vol. I, Section 1.10.6).

A.6 HEART

The user of this method is offered point values and uncertainty intervals of nominal error probabilities (NHEPs, in HEART denoted as "nominal human unreliabilities") for nine tasks described very generally. The point values are between 0.55 (task: "totally unfamiliar, performed at speed with no idea of the likely consequences") and 0.00002 (task: "respond correctly to system command even when there is an augmented supervisory system providing accurate interpretation of systems state"); the data are based on literature studies. From among these tasks one task should be selected which is applicable to the task to be analysed. HEART offers a list of 38 factors for the situation-specific modification (by multiplication) of the corresponding error probability. Each factor is described as an error producing condition (EPC). The numerical values specified for the factors range between $K = 17$ ("unfamiliarity with a situation which is important but which only occurs infrequently or which is novel") and $K = 1.02$ ("age of personnel performing perceptual tasks"). Most of these data are also based on literature studies, but literature references are missing for some factors.

The HEART user is now requested to select those EPCs which he considers relevant for the success or failure of the task investigated. For each selected EPC a value x (rating) defined in the interval (0;1) is to be estimated which describes the degree of the presence of the EPC. The multiplicative effect of an EPC thus evaluated is then quantified using the factor

$$[1 + (K - 1) \cdot x].$$

A.7 INTENT

INTENT is specialised in the quantification of decision-based errors due to errors of intention. This corresponds to the error types 3 to 6 outlined in Table 5-9. The INTENT user is offered a choice of 20 nominal errors. These errors were identified from a data bank with the aid of two computer programs and from North American nuclear operating experience. With respect to the causes that could lead to errors of intention, INTENT classifies these errors into four categories ([Gertman 92], Table 1):

1. Four errors due to the inclusion of possible action consequences, e.g. »tolerate an out of range situation with minor consequences«.
2. Four error attitudes leading to circumventions, e.g. »violate procedure and devise own«.
3. Six errors due to set crew response: e.g. »competing goal states lead to a wrong conclusion«
4. Six errors depending on internal (e.g. memory capacity) and external (e.g. written documents) resources (»resource dependencies«, e.g.: »crew consult inappropriate resource in emergency«.

For each error INTENT gives estimates for the lower bound (LB) and upper bound (UB) of the occurrence probability; these are expert estimates. Moreover, INTENT also includes a set of 11 (very generally and shortly described) performance shaping factors (PSFs) whose weighting factors (w_i , $i = 1 \dots 11$) were also determined by expert estimates.

The user of the INTENT method is now requested to select that error from among the 20 nominal errors which applies to the decision-based error for which a probability is to be estimated. Estimates for the 11 PSFs are then given on a 5-point scale. Taking the weighting factors into account it is then possible to calculate a reliability index SLI from these estimates, which is defined in the interval (0;1) ([Gertman 92], equation 2); this corresponds to the SLIM procedure. For the SLI-dependent calculation of the error probability p , however, a mathematical procedure deviating from SLIM is proposed. INTENT assumes that p is equal to the (1-SLI) 100 % quantile of a lognormal distribution with LB as the 5 % quantile and UB as the 95 % quantile.

$$\ln(p) = \sigma \cdot \phi^{-1}(1 - \text{SLI}) + \mu$$

with

$$\sigma = \frac{\ln(\text{UB} / \text{LB})}{3.29}$$

$$\mu = \frac{\ln(\text{UB} \cdot \text{LB})}{2}$$

where LB is the 5 % lower bound and UB the 95 % upper bound of the error probability.

APPENDIX B. TASK 94-1 SURVEY QUESTIONNAIRE

PWG 5/TASK 15 (Copy of Survey Questionnaire)

QUESTIONNAIRE ON CRITICAL OPERATOR ACTIONS

The focus of this questionnaire is on the treatment of Category C actions (see definition below). Section III, Treatment of Category C Actions, will form the bulk of the response. Section II, General Information, is requested in order to provide a context for understanding the results. This section and Section IV, Applications of HRA Results, are both expected to be relatively short. See also the attached example response.

I. ORGANIZATION REPLYING TO THE QUESTIONNAIRE

Name of Organization
Contact Person
Address and contact information [*address, telephone, fax, e-mail*]

II. GENERAL INFORMATION

1. PLANT CHARACTERISATION

Type [*BWR, PWR, or other (specify)*], Vendor, Architect Engineer, Site, Start of operation, Level of redundancy and physical separation, Backfits

2. PSA CHARACTERISATION

Level&Scope [*include modes of operation analyzed*], Year of publication, Originator, Language, Modelling Approach [*large ET/ small FT or small ET/large FT*], Availability for external users, Review status and availability, Core Damage Frequency [*mean value from internal events, external events, internal + external events analyses, as available*]

3. GENERAL CHARACTERISATION OF HRA

Overall approach and scope [*types of interactions considered/not considered, structure of the HRA analyses*], Plans for extensions (if any), Percent (rough) of the total effort [*man-months of total*]

4. PLANT-SPECIFIC FACTORS CONSIDERED IMPORTANT FOR HRA

Type of procedures [*symptom-based, event-based, with or without flowcharts, etc.*], Procedures considered/modeled in the HRA, 30-minute rule, Level of automation, Other

5. IMPORTANCE OF CATEGORY A AND B OPERATOR ACTIONS

Listing of specific actions if any identified among important contributors, Numerical perspective such as Fussell-Vesely importance or equivalent. *Similar information for Category C actions are to be included in Section II, Item 8, of this questionnaire.*

III. TREATMENT OF CATEGORY C ACTIONS AND RESULTS

6. METHODS USED IN THE TREATMENT OF CATEGORY C ACTIONS

Description including references and version numbers for each method (identification, definition, screening, quantification). Please be sure to include performance shaping factors or equivalent factors considered as well as data sources. Note that category C actions include recovery actions; descriptions of methods and criteria related to these should be included, e.g. criteria for crediting actions.

7. SCREENING APPROACH USED FOR CATEGORY C ACTIONS

Qualitative vs. quantitative, Probability level, List of actions considered, list of actions selected for detailed treatment. In some PSAs, a 'screening' approach may only be used for recovery actions (type 5); in others, no actions are credited for certain types of external events (e.g. seismic). Please elaborate.

8. RESULTS FOR CATEGORY C ACTIONS

8.1 IMPORTANT CATEGORY C ACTIONS

List the category C actions considered as important or dominant contributors to CDF and, if available, risk. A numerical perspective is important, include the available measures of importance [*contributor rank, Fussell-Vesely importance, split fraction importance, top event importance, etc.*]. Also include the mean probability value.

*In this questionnaire, **an operator action refers to each human interaction quantified separately.** The degree to which quantification is context or sequence specific will depend on the PSA; please provide the information that is available, indicating clearly whether the action refers to a specific context or sequence or to a class of actions (the same action in many or all sequences).*

Suggestion: as available, list those category C actions identified:

- a) as important contributors to the total (internal + external events) CDF,*
- b) as important contributors (actions) for the internal events-only PSA,*
- c) as important contributors (actions) for the external events PSA, and*
- d) (considered) most important in the HRA.*

8.2 DETAILED ACCOUNT OF IMPORTANT CATEGORY C ACTIONS (FROM 8.1)

For each of the cat. C actions listed in Sec. 8.1 (for at least the 5-10 most important actions), provide a detailed account addressing the following:

- Description of the action and its context (as a minimum, scenario or sequence and preceding events and actions)
- Important performance shaping factors for this action
- Other factors specific for this action that affect quantification e.g. dependencies
- Estimated probability mean value and other available distribution parameters. If possible and applicable, estimated values of the action ‘components’ [*as probability or percentage*], e.g. failure to initiate timely correct response and failure to execute required response correctly
- Sensitivity analysis results

8.3 PRINCIPAL PSA RESULTS RELATED TO CATEGORY C ACTIONS

Any further comments or insights regarding category C actions.

IV. APPLICATIONS OF HRA RESULTS

9. HRA-BASED IMPROVEMENTS OF DESIGN/PROCEDURES (IF ANY)

Description, Significance

10. USE OF PSA IN OPERATOR TRAINING

Description

11. Other

Description

-
- Category A (Type 1): interactions occurring prior to an initiating event. Typically related to test or maintenance
 - Category B (Type 2): interactions that lead to a plant transient (initiating event). In most PSAs, these are implicitly treated in the modeling of initiating events
 - Category C (includes Types 3, 4, 5): interactions taken by plant staff after the initiating event.

APPENDIX C. QUESTIONNAIRE ON DETAILED TREATMENTS

Copy of Detailed Treatment Questionnaire (last page has not been reproduced)

Detailed treatment for two actions Guideline for the additional contribution for ‘common actions’

BWR studies. 1) *manual depressurization* and 2) *standby liquid control (SLIC) actuation*

PWR studies. 1) “*Feed and Bleed*” and 2) either *Alignment for recirculation* or *Loss of RHR*

A. Concise description of action.

The items requested in Section 8.1 and 8.2 of the Task 15 Questionnaire, “Important Category C Actions” and “Detailed account of important category C actions” remain appropriate. See a suggested format for this part below.

B. Documentation of the quantification

For each action, describe how the human error probability was quantified. The objective is a “traceable” description that a reviewer can follow step by step, for example, from the PSF ratings to the HEP value. An example of a part B is provided (page 3); this documentation will of course depend on the method(s) used.

If you have any questions, please do not hesitate to contact Mr. Vinh Dang. Contact information:

Paul Scherrer Institute
CH-5232 Villigen PSI
Switzerland

vinh.dang@psi.ch
tel: +41 56 99 29 67
fax: +41 56 99 21 99

Suggested format for part A only, from Spanish contribution.

2) Action AH1FDYBL2FOIO (failure to RCS F&B)

- Initiating event: Generic Transient
- Initiating event frequency: 4.51 / year
- Initiating event contribution to the core melt frequency due to internal events: 7.46 %
- Sequence brief description, n° 7: Reactor trip, failure of secondary circuit heat removal, failure to RCS F&B.
- Sequence probability: 5.1 E-7
- Sequence contribution to the initiating event [*to the core melt frequency from the IE*]: 79.81 %
- Human action brief description: Operating team fails to detect or diagnose the necessity of RCS F&B or fails in the manual execution.
- Human error probability (distribution, mean value, error factor): lognormal, 8.4 E-3, 5
- Importance measures:

Fussell-Vesely: Below 2.99 E-2
 Risk increment: Below 5.17 E+1
 Risk reduction: Below 1.04

- Action objective: When Secondary is not available for removing heat from the RCS, the operating team should maintain the RCS coolant inventory and relieve the heat from the RCS to the Containment by using the HPSI and the Pressurizer relief valve.
 - Cognitive aspects: Operating team should detect that conditions for considering loss of SEcondary heat sink and for initiating F&B have been reached.
 - [• Notes concerning particularities of the plant design may also be provided if appropriate.]
 - Procedures: Some steps of the Critical Safety Function (Heat Sink) Recovery Guide
 - Critical actions: To initiate Safety Injection, to verify injection flow is adequate, to open both Pressurizer relief valves and to verify relief valves have opened.
 - Success criteria
 - Conditions for F&B are reached at: 1800 s
 - Available time: 300 s
 - Operating team response time: 50 s
 - Performance shaping factors considered
 - Operator experience: Average knowledge and training
 - Stress level: Grave emergency
 - Quality of operator/plant interface: Good
 - Cognitive processing type: Rule
 - Members of the operating team involved in the task: shift chief (SC), shift supervisor (SS), and reactor operator (RO).
 - Dependencies between operating team members
 - SC-RO: High dependency
 - SS-RO: Moderate dependency
- Non-response probability:
 Median probability: 7.0 E-3 (HCR)
- Manual error probability:
 Median probability: 1.45E-3 (THERP)

Example for part B only, from U.K. contribution, on next page, has not been included.

APPENDIX D. IMPORTANT ACTIONS (TABLES)

The important actions identified in the BWR and PWR PSAs are listed in tables in this appendix. This information is excerpted from the full questionnaire responses (Appendix F).

D.1 Important Actions for BWRs (by PSA)**Mühleberg (BWR)****Table 1. Important Cat. C Operator Actions**

Description	Importance	Mean Value
Failure of Manual Depressurization of Reactor Coolant System (RCS) Feedwater and RCIC Failed.	12.2 %	1.30 E-2
Manual depressurization of Reactor Coolant System (This class of actions includes the split fraction immed. above.)	14.5 %	1.30 E-2 – 3.97 E-2 (transients) 4.71 E-2 – 8.59 E-2 (ATWS)
Operator inhibits Automatic Depressurization System (ADS) and controls Low Pressure Injection (LPI) (ATWS)	1.52	1.18 E-1 – 1.32 E-1
Operator starts Shutdown and Torus Cooling System (STCS) in Torus Cooling (TC) mode	1.32	1.72 E-4 – 1.66 E-3 (transient/LOCA) 1.24 E-3 – 1.64 E-3 (ATWS)
Operator starts Standby Liquid Control System (SLCS) (ATWS)	1.04	1.47 E-2 – 1.77 E-2

TVO NPP (BWR)

Cat. C Action	F-V Importance	HEP
Connecting a DG supply from the neighboring unit	6.55 %	2.4e-1
Manual start of the boron supply	6.95	
Manual depressurization of the reactor	3.95	1.3e-2
Manual restart of the feedwater system	2.30	
Connecting electrical power supply from a hydropower station	1.76	
Manual restart of the residual heat removal chain	0.957	
Manual depressurization of the reactor with a relief valve stuck open	0.495	3.6e-2
Manual start of the refilling of the auxiliary feedwater tank	0.380	5.6e-3
Manual start of an alternate residual heat removal chain	0.307	
Backflushing of pump suction strainers of the containment spray system	0.252	

Both sets of values are from the Revised PSA.

- In the original PSA, one importance analysis for cat. C actions was performed by assessing the effect on CDF of assuming that a class of actions (same action in many sequences) always fails. In this analysis, manual start of the refilling of the aux. feedwater tank and manual depressurization of the reactor were the most important contributors to CDF.
- In the updated PSA, only Fussell-Vesely importance was calculated.

Dodewaard (BWR)

Cat. C Action (number indicates list order)	F-V Importance	F-V rank (among cat. C actions)	Risk Ach. Worth	R.A.W rank	HEP
3. Operator fails to inhibit ADS during ATWS	4.96 %	1	1.05	5	5e-1
5. Operator overfills RPV with ECCS during ATWS - ADS not inhibited	4.1	2	1.1	3	3e-1
4. Operator fails to start SLCS during ATWS sequences	2.37	3	1.77	2	2e-2
2. Operator fails to initiate ADS during ATWS	2.08	4	1.08	4	2e-1
1. Operator fails to open ADS valves during non-ATWS sequences	0.773	5	78.4	1	1e-4

For actions with importance measures, additional importance measures are available: Risk Reduction Worth = $1/(1-F.V.)$, Birnbaum Importance.

The Dodewaard actions, “Operator fails to start wetwell cooling (and RHR in a “CS” sequence / in an ATWS event)” are probably similar to the Mühleberg action, “Operator starts Shutdown and Torus Cooling System (STCS) in Torus Cooling (TC) mode”. In the Dodewaard, they are quantified as 1e-5 (5 hrs available) and 1e-2, respectively, but are not among the important contributors.

The following “were identified as Contributors to Safety in the Dodewaard Shutdown PSA”:	
Operator fails to isolate or maintain level upon low level alarm.	[No HEP/no importance]
Operator fails to maintain level or fails to refill.	
Operator fails to initiate Long Term Cooling given that RPV is open.	
Operator fails to open ADS valve.	
Operator fails to open Isolatin Condenser return valve primary side manually.	
Operator fails to start reactor water clean-up system.	
Operator fails to diagnose Loss of cooling and attempt to re-establish cooling.	

B1100 (BWR)

$$\text{Critical importance} = \text{HE}/\text{CDF} \cdot \partial \text{CDF}/\partial \text{HE}$$

Cat. C Action	Critical importance	HEP
Failure of making lineup of decay heat removal (DHR) using containment venting system in loss of residual heat removal (RHR) function	0.30	2.5e-2
Failure of the manual initiation of SLCS in ATWS	0.23	2.7e-1
Failure of the manual depressurization of the reactor pressure vessel using low pressure injection systems	0.11	2.9e-3

SBWR (BWR)

Cat. C Action	F-V Importance (Risk Red. W)	F-V rank (among cat. C actions)	Risk Ach. Worth	R.A.W rank	HEP
Failure to recognize the need or to actuate the ADS during SLOCA	30.5 % (1.439)	1	4.16	3	8.8e-2
Failure to recognize the need of low pressure makeup	21.9 (1.281)	2	5.74	1	4.4e-2
Failure to recognize or to check the level decreasing in the RPV	4.02 (1.042)	3	4.90	2	1.0e-2

D.2 Important Actions for PWRs (by PSA)**Alm (PWR)****Important Type 3 actions**

Description	F-V Importance (RRW)	Risk Red. Worth	Risk Increment Worth	Median HEP
Crew fails... ♣ also includes failure to execute the action within the available time and failure in manual execution.				
1. to line up CCW manually as required for supporting the recirculation phase of the LPSI (SLOCA)	6.78 %	1.07	83.8	5.08e-4
2. to detect or diagnose the necessity of changing from HPSI injection phase to recirculation phase or fail to execute it within the available time (MLOCA)	< 2.99	1.04	51.7	4.96e-4
3. to line up manually the valves between the LPSI pumps and the RCS through the HPSI pumps for initiating the recirculation phase of the HPSI (MLOCA)	3.68	1.04	< 51.7	8.28e-4
4. to control the secondary circuit heat removal manually using the motor-driven pumps of the Auxiliary Feedwater System (AFW) (generic transient)	3.22	< 1.04	168	1.2e-4
5. to detect or diagnose the necessity of RCS F&B. ♣ (generic Transient)	< 2.99	< 1.04	< 51.7	8.4e-3
6. to detect or diagnose the necessity of starting train B of CCW and Service Water System (SW) when the train A of the Ventilation Unit for the AFW motor-driven pumps area has failed. ♣ (generic Transient)	< 2.99	< 1.04	< 51.7	6.97e-3
7. to identify or isolate the affected SG (SGTR). ♣	< 2.99	< 1.04	< 51.7	3.15e-4
8. to detect or diagnose the necessity of starting the RCS depressurization down to RHR conditions below low level in the RWST is reached. ♣ (SGTR)	4.81	1.05	< 51.7	3.97e-3
9. to detect or diagnose the necessity of rerunning the equipments necessary for coping with the (total loss of outside electrical power). ♣	3.05	< 1.04	< 51.7	7.69e-2

See separate table for recovery actions (no importance measures)

Alm (PWR) (continued)

Some Type 5 actions (recoveries)*

Description	HEP
Open interconnect valve by bypassing interlock to permit the recirculation phase of the HPSI. (interlocked with failed or miscalibrated pressure transmitter) (MLOCA)	3.95e-2
Open manually (locally) the discharge valve of the turbine-driven pump of the AFW closed during testing. (Generic transient)	9.0e-3
Open RHR suction line valves by bypassing interlock (interlocked with failed or miscalibrated RCS pressure transmitter) (SGTR)	4.2e-3
Line up the standby pump (to recover failure of the three CCW pumps) (Loss of CCW)	4.2e-3

* “These actions recover specific failures in specific minimal cut sets... so the importance measures from the point of view of total CMF are not available with these kinds of descriptions limited to one specific sequence.”

AP600 (PWR)

Table. Most Important Cat. C Operator Actions

Description	Risk Reduction Worth	F-V Importance (calc)	Risk Ach. Worth	Mean HEP
Failure to depressurize the RPV following events with failure of the high pressure systems	1.0	< 0.1 %	11.8	5.3e-4
Failure to recognize the need of low pressure injection initiation following failure of high pressure systems	1.02	2.	10.4	2.2e-3
Failure to trip the reactor by deenergizing the MG sets following ATWS	1.10	9.	5.7	1.5e-2

Beznau (PWR)

Table. Ten Most Important Independent Split Fractions Related to Operator Actions

Description	Split Fraction Importance	Split Fraction Value (Mean Failure Prob.)
Initiate Spray Recirculation After SGTR or PTS	7.74 %	5.16e-1
Initiate Spray Recirculation After Core Melt	6.62	2.51e-1
Operator Depressurizes Plant (SGTR) - No other Failures	3.20	2.26e-3
Bleed and Feed Cooling - Severe Over Cooling Case	2.60	1.56e-1
Maintain Feedwater Flow and Heat Removal from at least 1 SG (General Transient, Steam Line Break Outside Containment)	2.60	1.20e-1
Initiate Spray Recirculation After LLOCA	2.39	6.43e-1
Operator Control SI (HPI Pumps) for PTS	2.35	9.00e-1
Bleed and Feed Cooling - Given LBW Failed	2.19	1.59e-1
Bleed and Feed Cooling - Moderate Over Cooling	1.78	1.56e-1
Operator Depressurizes Plant (SGTR) - DC and MS Failed or UO Failed ^b	1.62	2.56e-2

^b DC= 3/3 Downstream ARVs close after steam relief; MS= MSIV isolation failure; UO= Steam relief via upstream ARVs

For Beznau, Cooldown to Allow LPSI (CD for LPSI) is modeled within the top event OA (which includes both cooldown actions and feed and bleed actions). This top event (class of actions) has an importance of only 0.0887, of which approx. 96% is due to dependencies.

Borssele (PWR)

10 Most important "Human Errors" in the list of the top 75 most important events
ranked by F-V/Risk Reduction Worth

Cat. C Action C1 - manual backup C2 - EOP based C3 - recovery actions Operator fails to...	F-V Imp. Risk Red. Worth	Risk Ach. Worth	F-V or RRW rank	RAW rank	cut set of top 20	HEP pt. est.
override low seal water pressure transmitter trip (C3) (very sm. LOCAs) TJ-TN-OVRD-SY-HE	14.4e-2 1.168	3.00	3 (1)	89 (4)	1 16	6.7e-2
throttle LP-ECCS pump to prevent runout (C2) (LLOCA) TJL-RUNOUT-SY-HE	4.89e-2 1.051		18 (2)		6	1.2e-2
open PORVs adequately for F&B (C2) (SLOCA) YJ-PORVSFB-SY-HE	4.31e-2 1.045		19 (3)		17	4e-1
manually borate the primary with the max. concentration (C2) TB-TKBORAT-SY-HE	2.92e-2 1.03		23 (4)			6.2e-2
isolate flow div. through failed accum. check valve (C3) TJ-DIVRECV-SY-HE	2.29e-2 1.023		27 (5)			6.1e-1
manually initiate TJR (sump recirc. mode of ECCS) system during Large LOCA (C1) (LLOCA) TJ-TJRACTU-A-HE	2.06e-2 1.021		35 (6)		12	1e-0
Trip one operating pump of CC system or isolate 1 HX of aux & emergency cooling water system (C3) VF-ISOL-PM-SY-HE	1.30e-2 1.013		43 (7)			6e-1
initiate 100 °C/hr cooldown during LOCA, 1 switch in control room (C2) (SLOCA) SCD-3HOURS-SY-HE	1.23e-2 1.012	41.84	44 (8)	27 (3)	17	3e-4
initiate normal cooldown, one switch operation from control room (C2) (very sm. LOCA) SCD-12HOUR-SY-HE	1.14e-2 1.012	104.32	45 (9)	15 (2)	18	1.1e-4
control HP-ECCS flow within 10 minutes TJH-CTLREC-10-HE	1.09e-2 1.011		46 (10)			5e-1
manually start TJR (RHR) heat removal system (C2) TJ-RHRACTU-SY-HE		127.81		12 (1)		n/a

Borssele (PWR) (continued)

Eight actions appear in the top 100 most important events, when these are ranked by Risk Achievement Worth. Some of these overlap with the human actions from the F-V ranking..

	F-V Imp. Risk Reduct. Worth	Risk Ach. Worth	F-V or RRW rank	RA W rank	HEP point estimate
<i>TJ-SENSTRX-SY-HE</i> <i>Dependent miscal. of accum. pressure</i> <i>sensor/transmitters (A)</i>	<i>4.14e-2</i> <i>1.043</i>		<i>20</i>		<i>1e-2</i>
<i>YZ-TJ-L051-LT-HE</i> <i>CC miscal. of inundation tank level</i> <i>transmitters (A)</i>	<i>2.79e-2</i> <i>1.029</i>	<i>6.55</i>	<i>24</i>	<i>54</i>	<i>5e-3</i>
<i>YZ-YA00PCC-PT-HE</i> <i>CC miscal. of YZ-101 press. transmitters</i> <i>primary system (A)</i>	<i>2.27e-2</i> <i>1.023</i>	<i>5.52</i>	<i>28</i>	<i>76</i>	<i>5e-3</i>
<i>TJ21-24TESTSW-HE</i> <i>Improper restoration of priority ('Vorwahl')</i> <i>switch following flow test of LP-ECCS Pump</i> <i>(A)</i>	<i>2.06e-2</i> <i>1.023</i>	<i>3.24</i>	<i>29</i>	<i>82</i>	<i>1e-2</i>
<i>YZ-YA0P-36-PT-HE</i> <i>CC miscal. of YA01P53/63 Comparators -</i> <i>YZ36 (A)</i>		<i>3.39</i>		<i>76</i>	

Doel (PWR)

Category A Important Actions

	<i>Action</i>	Global Human Error Probability	% of <i>TCDF</i>
	Position error of a manual valve used to test the ultimate injection system (RJ) to the seals of the main coolant pumps (MCPs). This system is 3-monthly tested.	$1,5 \cdot 10^{-3}$	
#	<i>Sequence description</i>		
①	Loss Of Off-Site Power (LOOP) inducing a Seal LOCA with failure of HPSI		0,6 %
②	Loss Of Electric Power (LOEP) inducing a Seal LOCA with failure of PORV, LPSI injection or recirculation		0,4 %
③	General Plant Transient (GPT) inducing a Seal LOCA with failure of HPSI		0,3 %
	All MCS including this HE		2.0 %
	Position error of the non-redundant valve on the ultimate seal injection system (RJ) line (monthly tested) Recovery factor : periodic alignment control with written check list.	$1,5 \cdot 10^{-3}$	
	Position error of valve on tank suction line of ultimate seal injection (3-monthly tested). Recovery factor : periodic alignment control with written check list.	$1,5 \cdot 10^{-3}$	
#	<i>Sequence description</i>		
①	Loss Of Off-Site Power (LOOP) inducing a Seal LOCA with failure of HPSI		0,3 %
②	Loss Of Electric Power (LOEP) inducing a Seal LOCA with failure of PORV, LPSI injection or recirculation		0,2 %
③	General Plant Transient (GPT) inducing a Seal LOCA with failure of HPSI		0,2 %
	All MCS including this HE		1,2 %

Doel (PWR) (continued)
Category C Important Actions

Identifier	Action Description	Global Human Error Probability	% of TCDF Modes 1-3	% of TCDF Modes 4-7	% of TCDF all modes
E2-E8 (all sequences)	Operator fails to manually start in due time one stand-by CVCS pump combined with the failure on demand of ultimate system RJ following the loss of the electrical 110V bus.	$1.1 \cdot 10^{-1}$	10.0	-	9.5
E0-E2 (all sequences)	Operator fails to recover the loss of the 110V electrical bus.	$3.6 \cdot 10^{-2}$	7.6	-	7.1
Z-E2 (all sequences)	Operator fails to locally open valves in an acceptable time delay for SI pump start-up in modes 5, 6 & 7.	$5.4 \cdot 10^{-2}$	-	34.4	2.6
U-E4	Operator fails to secure LPSI pumps before manually switching over to recirculation.	1	-	0.2	0.02
HF NO U2	No application of the U2 procedure (primary feed & bleed), resulting from failure to diagnose the U2 procedure in case of total loss of FeedWater to SGs (Auxiliary FW, Emergency FW and Main FW failure).	$1.8 \cdot 10^{-3}$	1.7	-	1.6

DRS (PWR)

Action	Importance	Prob.
Shutdown with 100°C/hr with a wrong gradient within 30 mins after accident		1e-2 (median)
No shutdown with the emergency system in 2-3 hrs after accident (small leakage at the pressurizer)		1e-3 (median)
Primary Feed and Bleed		0.1-1 (mean)
Secondary Feed and Bleed		0.01-0.1 (mean)

Loviisa (PWR)

Operator action failure (model may include hardware failure*)	Identifier	Importance (F-V)	Comment
operator fails to recover UV25/20 (instrumentation room ventilation)	REC25/20K	5.48E-2	Class of actions
operator fails to feed steam generators during LIRV	LIRVFEEDQH	5.34E-2	Specific context
failure to start additional emergency feed water system	SGF00000NH	5.08E-2	Class of actions
operator fails to isolate VLOCA leakage in 2 hours	OPERCOLDQH	3.84E-2	Class of actions
operator fails to provide RCPs integrity	LIRVRCPSQH	2.71E-2	Specific context
recovery, operator fails to recover loop washing	RECWASHOPE	2.05E-2	Specific context
operator fails to stop TK/TB pumps (volume and chemical control system pumps)	RECOVERYQK	1.19E-2	Specific context
recovery, seal injection from TK or VF13- seal cooling fails	RECSEALINJ	8.55E-3	Class of actions
operator fails to recover safety injection sump operation	SUMPRECONH	4.10E-3	Class of actions
operator fails to isolate VLOCA leakage in 0.5 hours	OPERAHOTQH	3.9E-3	Specific context

* Many of these include not only operator actions but also component failures. The latter have been considered minor compared to the human errors; FV still gives a good indication about the operator action importance.

P1100 (PWR)

Table. Six Most Important Cat. C Operator Actions

Description	F-V Importance	Mean HEP
F&B operation with preceding operator failure to isolate a faulted SG in SGTR	7. %	1.4e-2
F&B operation in SLOCA and transients	6.	3.9e-3, 4.2e-3
F&B operation with preceding operator failure to isolate aux. feedwater in secondary side break	6.	4.2e-3
Isolation of a faulted SG in SGTR	5.7	4.2e-3
Isolation of CCWS in Loss of Component Cooling Water System	5.7	7.6e-2
Isolation of ISLOCA	3.7	0.3

P900 (PWR)

The importance measure is “contribution to overall core melt frequency” [F-V].

Action (classes)	Importance	Prob.
Water make-up under loss of cooling during shutdown (RHRS failure / LOCA / dilution)	17.6 %	1– 5e-3
Protection of CVCS pumps (total loss of terminal heat sink)	10.6	1.1e-1
Back-up of LPSI or CSS pumps by mobile device (Long term post-LOCA situation -Loss of LPSI or CSS)	6.2	
Feed and Bleed (total loss of S.G. feedwater)	2.6	
Isolation of failed SG and depressurization (SGTR)	2.4	
Subcriticality (manual scram - boration) in ATWS	1.9	
Rapid cooling by SG to reach LPSI actuation (small LOCA - HPSI failure)	1.7	

P1300 (PWR)

The importance measure is “contribution to core melt frequency”, which is sum of probabilities of the sequences with the action divided by CMF.

Action (classes)	Importance (power POSs)	Importance (shutdown POSs)	Importance total CMF	HEPs (selected values)
Water make-up (loss of RHR, various breaks)	0.06 %	61. %	61. %	1.12e-3
Feed and bleed	7.9	0.24	8.2	5.55e-3, 1e-1
Spurious shutdown of systems (various breaks)	4.1	0.04	4.2	1.2e-3, 4e-3
Interruption of dilution	0.9	4.3	5.2	
Sub-critical state after ATWS	1.3	–	1.3	
Rapid cooling (failure of rapid cooling) (intermediate break)	0.9	0.12	1.0	

- 11 of top 20 shutdown sequences, including top 8, involve failure to provide make-up water.
- 70% of feed and bleed failure, which is modeled as the failure of procedure H2 for the total loss of feed, is due to the operators’ (and the safety engineer’s) failure to enter the feed and bleed mode.

Sizewell (PWR)

Action	Importance	Prob.
Bleed and Feed within approx. 1 hr of reactor trip followed by loss of all feed	1.591e-1	1.1e-1
Operator starts motor-driven auxiliary feedwater pumps following loss of the primary low voltage essential electrical system	1.159e-1	1.0e-1
Bleed and Feed within several hrs of reactor trip followed by loss of all feed	1.610e-2	3.5e-2
Operator actions following loss of RHR cooling under situations of reduced primary circuit inventory (Failure to diagnose loss of cooling)	1.528e-2	5.0e-5
Manual control of battery charging supplies following simultaneous loss of all 11 kV and of the corresponding emergency diesel generator (Fail to start Battery Charging Diesel)	9.874e-3	3.4e-1
Operator actions following loss of RHR cooling under situations of reduced primary circuit inventory (Failure to use running CVCS pump)	9.167e-3	3.0e-5
Manual isolation of leaks (via the RCP thermal barrier or via the RHRs heat exchanger) into the CCWS (Isolation within 1 hr)	8.211e-3	5.5/5.3e-2
Realignment of the auxiliary feedwater to the town's water reservoir	9.011e-3	2.6e-4
Manual control of battery charging supplies following simultaneous loss of all 11kV and of the corresponding emergency diesel generator (Fail to switch over battery chargers to secondary group boards)	7.889e-3	2.7e-2
Operator actions following loss of RHR cooling under situations of reduced primary circuit inventory (Failure to use stand-by High Head Safety Injection (HHSI) pump)	5.559e-3	2.6e-3

APPENDIX E. THE ATHEANA METHODOLOGY: EXTENDED SUMMARY

I. Introduction

This appendix presents an outline of an analytical process for performing a human reliability analysis (HRA) in the context of a probabilistic risk assessment (PRA), that addresses the major deficiencies of current HRA methods, and, in particular, provides an approach to the analysis of errors of commission. This analytical process has been developed using the concepts captured in the multidisciplinary framework described in NUREG/CR-6265 [Ref. 1], and supplemented with the experience obtained from the analysis of historical events at low power and shutdown, as described in NUREG/CR-6093. Both of these documents are earlier products of the project initiated by U.S. Nuclear Regulatory Commission in response to the recognized need for an improved, more realistic, approach to the modeling of human-system interactions. The framework recognizes the need to bring together the disciplines of behavioral science, cognitive psychology, and systems analysis, as well as input from plant operations, in order to capture realistically the human-systems interactions and their impact on safety. The analytical process is the application phase of a new approach to human reliability analysis. This approach, called ATHEANA [Ref. 3] (A Technique for Human Error Analysis), is based on an understanding of why human-system interaction failures occur, rather than on a behavioral, phenomenological description of operator responses, and represents a fundamental change in the approach to human reliability analysis. Section II presents an overview of the ATHEANA method, and Section III describes the application process.

II. Overview of the ATHEANA method

There are important human performance issues which are addressed in the ATHEANA HRA method to make the required improvements in HRA/PRA applications. The issues which represent the largest departures from current HRA methods all stem from the need to better predict and reflect the "real world" nature of failures in human-system interactions, as illustrated by past operational events. Real operational events frequently include post-accident errors of commission, which are minimally addressed in current HRA/PRA. The occurrence of an error of commission is strongly influenced by the specific context of the event (i.e., plant conditions and performance shaping factors). This specific context of an event frequently departs from the nominal plant conditions assumed by PRA and HRA analysts to represent the plant conditions during off-normal incidents.

Consequently, the HRA modeling approach adopted for ATHEANA is a significant shift from current approaches. In particular, to be consistent with operational experience, the fundamental premise of ATHEANA is that significant post-accident human failure events, especially errors of commission, represent situations in which the context of an event (plant conditions, PSFs) virtually forces operators to fail. It is this focus on the error-forcing context which distinguishes ATHEANA from other HRA methods.

The ATHEANA modeling approach involves more than simply a new quantification model. Included in ATHEANA is a better, more comprehensive approach to the identification and definition of appropriate human failure events (HFEs), and the placement of these human failure events in the PRA model. The

guidance on how to search for the HFEs is based on an understanding of the causes of human failures as indicated above.

In applying ATHEANA to a PRA, the representation of post-accident HFEs which are errors of commission will be similar to the representation of errors of omission already addressed by existing HRA methods, in that they will be identified and defined in terms of failure modes of plant functions, systems, or components. However, errors of omission (EOOs) result from failures of manual operator actions to initiate or change the state of plant equipment. Therefore, EOO definitions typically are phrased as "operator fails to start pumps", for example. Errors of commission, on the other hand, result from specific actions on the part of the operators. Generally, post-accident errors of commission result from one of the following ways by which operators fail plant functions, systems, or components:

- by turning off running equipment;
- by bypassing signals for automatically starting equipment;
- by changing the plant configuration such that interlocks that are designed to prevent equipment damage are defeated; and
- by excessive depletion or diversion of plant resources (e.g., water sources).

In a PRA model, only the most significant and most likely HFEs need be included. Identification of the most likely is based on an understanding of the causes of error.

An HFE may result from one of several unsafe actions'. Application of ATHEANA involves, for each HFE, the identification and definition of unsafe actions and associated error-forcing contexts (EFCs). The identified error-forcing contexts (i.e. , plant conditions and associated PSFs), and their underlying error mechanisms, are the means of characterizing the causes of human failures. An unsafe action could be the result of one of several different causes.

Implicit in the definition of the HFEs and unsafe actions is the recognition that, because of the nature of nuclear power plant operational characteristics, there is generally time for the operators to monitor the changes they have initiated, which allows them opportunities to recognize and correct errors. Thus, the unsafe action is a result of an error and a failure to correct that error before the failure associated with the PRA basic event occurs. Therefore, the error forcing context associated with an unsafe action must address the factors that impact both the initial error and the failure to recover.

In the application of ATHEANA, the prioritization of HFEs will be based on the probabilities of the contributing unsafe actions, and these in turn will be based on the probabilities of the associated EFCs. Quantification of the probabilities of corresponding HFEs will be based upon estimates of how likely or frequently the plant conditions and PSFs which comprise the error-forcing contexts occur, rather than upon assumptions of randomly occurring human failures. Therefore, quantification of an HFE using ATHEANA is based upon an understanding of the following:

- what unsafe action(s) can result in the HFE whose probability is being quantified?
- what error-forcing context(s) can result in the unsafe action(s) comprising the HFE?
- how likely are these error-forcing contexts to occur?

As discussed above, there are two sets of EFC elements to consider: those associated with the initial error, and those that impact the potential for recovery. There may be common EFC elements between the two sets, and therefore the EFCs for a given unsafe action will be given by the union of the two sets of elements.

ATHEANA will be supported by two documents. The first, called the Frame of Reference Manual contains the knowledge required to apply the method. A related paper to be presented at PSA 96 describes the Frame of Reference Manual'. The second is called The Implementation Guidelines, and describes how to apply this knowledge in a plant specific manner. The underlying analytical process for application of ATHEANA is described below.

III. The ATHEANA process

The ATHEANA application process has been discussed in detail in Reference 5. Since publication of that document, considerable progress has been made on the development of the Frame of Reference Manual, and the discussion of the process presented here reflects this progress. The general structure of the process, captured in Figure 1, "Flow diagram of analytical procedure", taken from Reference 5, is still valid. Five tasks are identified. These are:

1. Identification of the candidate human failure events to be modeled;
2. Identification of potentially important types of unsafe actions that could cause each HFE;
3. For each type of unsafe action, identification of the most significant reasons for that type of unsafe action to occur, and for each type of unsafe action and its associated reason, identification of the potentially significant error-forcing contexts;
4. For each type of unsafe action and its associated reason, estimate the likelihoods of the error-forcing contexts and the consequential probabilities of the unsafe actions; and
5. For each HFE, sum the likelihood's of the error-forcing contexts and consequential probabilities of the unsafe actions for all potentially important types of unsafe actions that could cause the HFE.

Presented in this way, the process appears to be somewhat open-ended, since there is potentially a very large number of combinations of HFEs, unsafe actions, and error-forcing contexts that could contribute to the occurrence of a severe accident. However, the ATHEANA method Implementation Guidelines will incorporate detailed guidance on how to prioritize and/or screen the HFEs and unsafe actions for the most significant. At the highest level, the following general prioritization criterion is proposed:

- The unsafe actions of most interest are those that are taken on a rational (if incorrect) basis; that is, irrational, spontaneous, and arbitrary actions are not considered.

This section discusses how the information presented in the Frame of Reference Manual will be used in the five tasks identified above and in Figure 1, and how the high level criterion is implemented.

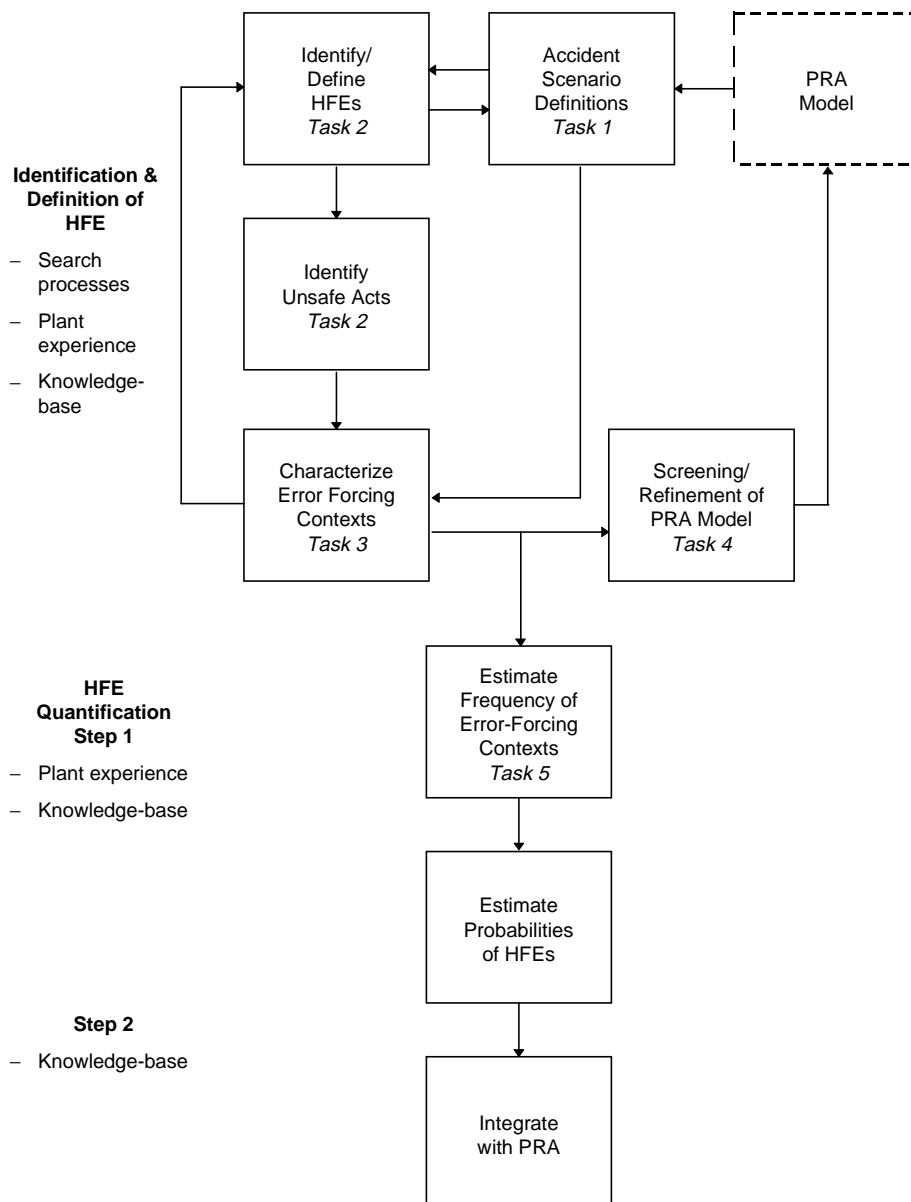


Figure E-1. The ATHEANA process flow diagram [adapted from Parry et al., 1995]

Task 1: Familiarization with PRA Model and Accident Scenarios

It is assumed here that the analysis starts with an existing PRA. To analyze the human-system interactions within the context of that model it is necessary to become familiar with the definitions of all the elements of that model, and with the accident scenarios identified. It is also essential to understand the assumptions underlying the PRA model. Clearly, a key element in performing an HRA is to identify the role of the operating crew in mitigating, or controlling the progress of the accidents represented by the scenarios. It is essential to become familiar with the set of applicable procedures, and to understand which, and under what plant conditions, procedures are required. It is also necessary for the human reliability analyst to develop a clear picture of how the plant responds to the functional failures and operator actions represented in the scenarios.

Task 2: Identification of Potential Human Failure Events and Associated Unsafe Actions

This task is performed in two stages. The first is to identify the potential human failure events (HFEs). The second is to identify the potential human unsafe actions.

Identification of Human Failure Events

The selection of individual HFEs is based on the system or functional requirements associated with the events associated with the event tree branch points. The HFEs will correspond to human caused failures at the function, system or component level, and at this stage are defined entirely in terms of failure modes. (Note that, as discussed in Reference 5, the events that are finally incorporated in the PRA model may be defined more precisely as arising from specific unsafe actions and specific reasons.) Since the failure modes in terms of their impact on the system are similar to those caused by hardware faults, they are limited in number.

For example, for the initiating event "loss of main feedwater" in a PWR, the preferred method of decay heat removal is generally identified as the use of the auxiliary feedwater system. The auxiliary feedwater system is a standby system for which the success criterion in this scenario is, for example, that one out of three pumps start automatically and that the system continues to provide water to the steam generators for 24 hours following loss of main feedwater. The analysis should consider the following HFEs as causes of failure of the auxiliary feedwater system. They are identified as resulting from errors of commission (EOC) or errors of omission (EOO) to highlight the types of HFEs that are not normally included in PRAs.

Auxiliary feedwater system (AFWS) required to start on demand:

AFW equipment removed from automatic control (EOC), Automatic start of AFWS not backed up when required (EOO).

AFWS required to continue running:

Emergency operating procedures, (EOPs) require that manual control of the AFW system is established following initiation. The appropriate human failure events are:
AFW resources inappropriately diverted (EOC or EOO), AFW resources inappropriately depleted (EOC or EOO), Operating AFW equipment inappropriately terminated (EOC), Operating AFW equipment inappropriately isolated (EOC), Equipment operation results in under-feeding/filling (EOC or EOO).

The errors of commission have not generally been represented in existing PRA models. Typically, only errors of omission, such as failure to start the AFWS manually as a backup to the auto-initiation signal, or failure to make up to the CST to supplement the inventory, have been included. These new HFEs should be considered as new failure modes for the top event Auxiliary Feed Water System failure.

The Identification of Potential Unsafe Actions

The next step in the process is to identify the different ways in which the operators could produce the effect characterized by the failure mode identified above. This requires a detailed understanding of the systems and how they are operated. The Frame of Reference Manual will contain guidance on the generic types of unsafe actions that might occur. These can be specialized to the AFW system, by identifying those that are applicable to that particular system. This requires detailed knowledge of system design and

operational characteristics. The following are examples of specific unsafe actions that could apply to the AFW system:

Errors of Commission

The HFE 'AFW equipment removed from automatic control', could result from the following:
 Initiation signals bypassed or suppressed,
 Automatic signals taken out of "armed" status by placing pump start switches to manual,
 Motive and/or control power to the pumps removed or disabled,
 Taken out of standby status (e.g., pumps in "pull-to-lock").

The HFE 'AFW resources inappropriately depleted', could result from:
 CST inventory being depleted prior to equipment initiation,
 AFW equipment not re-aligned to secondary source when CST depleted.

Errors of Omission

The HFE 'Automatic start of AFWS not backed up when required', is already at the level of an unsafe action, and no further decomposition is needed.

Task 3: Identification of the Most Probable/Significant Causes of the Unsafe Actions (EFC)

The purpose of this task in the application process is the identification and prioritization of the EFCs that are associated with the unsafe actions. This is the critical task that makes ATHEANA different from all other HRA methods. Essentially what the task entails is the construction of models for the causes of the unsafe action in terms of failure modes of the activities identified in a model of information processing (see for example Reference 6), and the EFC elements associated with those activities. For example, one possible model may begin by exploring the initial problem as being a failure in situation assessment. This results in an incorrect situation model, which leads to an inappropriate response plan, which if carried out correctly results in the unsafe action. However, the model for the unsafe action must also take into account the failures of the operators to realize their situation model is incorrect and take corrective action. The opportunity for this to occur may be after the response has been executed and the operators are monitoring the plant to determine whether the effect of the actions they have taken are having the expected effect. The opportunity may, on the other hand, occur before the response execution has been completed, and could be triggered by new information.

In accordance with the analysis criterion that the actions of interest are those in which the crew behave in a rational manner, it is assumed that the operators are responding in accordance with "rules", which could be formal, e.g. , procedures, or informal, e.g., good practice. The method of analysis then considers the identification of the rule that, when inappropriately applied, results in the unsafe action, and identifying the reasons why that rule could have been invoked. A similar approach was adopted in Reference 7, but in that work, only the formal rules provided by the emergency operating procedures were investigated. Furthermore, the mode) of causes of unsafe actions adopted was crude compared with that developed for ATHEANA.

Building the models of causes of unsafe actions requires making use of several different types of information. It uses information that characterizes how errors can occur in the different stages of information processing, and the factors that influence the occurrence of errors. In addition, it is necessary to understand how information can be distorted by plant conditions and design features so that operators

can become confused as to the interpretation of indications. Generic descriptions of the ways in which plant physics/behavior, the algorithms that are used in instruments/indications, and other plant conditions can create confusion will be presented in the Frame of Reference Manual.

In the following paragraphs, a systematic and efficient approach to identifying the EFCs is outlined, focusing on the failures caused by problems in situation assessment. The approach is presented as a number of steps. At a high level, the unsafe action is considered to have arisen because the operators have an incorrect situation assessment model and fail to update it in a timely manner.

The first step is to determine if there is a rational explanation of why the unsafe action could be committed. This is done by identifying rules that the operators might apply to justify their actions, and which could apply for the PRA scenario of interest.

Step 1: For Each Unsafe Action Examine the "Rules " that Would Lead to the Unsafe Action

This is the first of a set of screening steps, and it is justified by the requirement that the operators' actions be rational. The purpose of this step is to identify reasons why the unsafe action would be performed, in terms of formal or informal rules of operation. So for example, for the unsafe action "SI inappropriately secured", the following types of rules might apply:

Formal

Procedure ES 0. I, Step x, SI Termination Criteria

Informal

The informal rules relate to behavioral responses that are ingrained as a result of training, such as:

- Avoid going solid in the pressurizer,
- Stop spurious SI,
- Protect pump when you get a trouble alarm.

The next step is to determine what information the operators would use to apply the rules, and where the information would come from, and is essentially an information gathering step.

Step 2: Identify Information Needed to Use the Rules

This step should identify both the primary and secondary sources of plant information that might be used, and the standard practices that are adopted, e.g. , look at ammeters as well as pump indicators. This can be regarded as establishing the operational practice that apply. Examples of the information needed to determine whether the conditions for applying the rule for SI termination are satisfied are:

Formal Rule (as identified above):

- Pressurizer pressure
- Pressurizer level
- Subcooling margin
- Secondary heat sink

- AFW flow
- Steam generator level

Informal rule:

Pressurizer level (if pressurizer solid rule)
 Pressurizer level and pressure (if spurious SI rule)

The next step is to identify the ways in which the criteria in the rules could have been interpreted as having been met, even though they have in fact not been met. Essentially this requires that the information that is available has to have been distorted, either by plant conditions, or by operator bias, or indeed both.

Step 3: Determine how the Rules Could Appear to Have Been Met When They in Fact Have Not Been

This step identifies the ways in which the incorrect situation model could have arisen, and also how information retrieval problems, plant conditions and physics problems, and operator problems (e.g., wrong mental model) could distort the information that should be seen by the operators.

However, in accordance with the PRA defined unsafe action, the incorrect situation model has to exist and persist until the unsafe action has manifested itself as a failure. This is addressed in the next step.

Step 4: Determine how the Operators Could Fail to Recognize that the Situation Model is Incorrect, and Correct it to Prevent Incorrect Application of Rule

This step is associated with the potential for recovery, and is needed to address the dynamic nature of response. This consists of a consideration of all stages of information processing. In this case, however, when the operators have initiated a response, the monitoring of the plant response, to confirm the appropriateness of the response that has been implemented is a key element for analysis.

Step 5: Identify the Potentially Significant Error-Forcing Contexts

This step corresponds to summarizing and analysing the information obtained in the previous four steps. A key point here is to search for the EFC elements that both cause the initial error and inhibit the possibility of recovery. Selecting the potentially most significant depends on several factors, which are plant-specific, and will require the input of plant experts. That the EFCs are expected to contain several elements is illustrated by the EFC identified for the following real event.

Crystal River 3 stuck-open pressurizer spray valve:

In this event the unsafe action was an inappropriate termination of High Pressure Injection, and the following EFC elements contributed:

- pressurizer spray valve position indication was inconsistent with actual valve position (due to pre-existing hardware failure and design);
- no direct indication was available of pressurizer spray flow;
- evolution in progress was to increase reactor power (basis for the erroneous conjecture that under-power event occurred).

Task 4: Refinement of HFE Definitions and Integration into PRA Logic Model

The issues associated with the refinement of HFE definitions are discussed in Reference 5 and are not reproduced here. This is primarily a systems analysis function to address the potential for dependencies. The Implementation Guidelines will give guidance on how to account for dependency given the identification of the most significant EFCs associated with the unsafe actions and HFEs.

Task 5: Estimate the Likelihoods of the Error-Forcing Contexts and the Consequential Probabilities of the Unsafe Actions

The approach to quantification is to first estimate the likelihood of the EFCs associated with each unsafe action, and for each EFC, to estimate the conditional probability of error.

- The estimation of likelihoods of EFCs will be based on constructing probabilistic models for the joint occurrence of the elements of the EFC. This entails obtaining estimates of the relative frequency with which the conditions in the EFC occur in the PRA scenario definition. That this is feasible has been demonstrated by the trial application in NUREG/CR-6350³.
- Estimation of conditional probabilities of error given an EFC. This will be a function of how specific the EFC definitions will be. In the case that the EFC is defined such that failure is almost guaranteed, then there is no need to estimate this conditional probability. However, in many cases, the EFC creates an environment in which the likelihood of failure is enhanced, and in this case, it will be necessary to develop methods for estimating these probabilities.

Once the HFEs to be included in the PRA model have been defined and incorporated into the logic structure in the appropriate way, the requantification of the PRA model is essentially trivial.

IV. Summary

This section has presented an overview of the ATHEANA method for HRA. The analytical process for application has been described, and the relationship to the two major documents that are in development, namely the Frame of Reference Manual and the Implementation Guidelines, has been discussed.

V. References

- /1/ Barriere, M.T., W.J. Lucas, Jr., J. Wreathall, S.E. Cooper, D.C. Bley, and A. Ramey-Smith, *"Multidisciplinary Framework for Analysing Errors of Commission and Dependencies in Human Reliability Analysis,"* NUREG/CR-6265, BNL-NUREG-52431, August 1995.
- /2/ Barriere, M.T. , Lucas, W.J. , Whitehead, D.W. , and Ramey-Smith, A. , *An Analysis of Operational Experience During LP&S and A Plan for Addressing Human Reliability Assessment Issues,* NUREG/CR-6093, Brookhaven National Laboratory: Upton, NY and Sandia National Laboratories: Albuquerque, NM, 1994.
- /3/ Cooper, S.E., Ramey-Smith, A., Wreathall, J., Parry, G.W., Bley, D.E., Taylor, J.H., and Lucas, W.J., *A Technique for Human Error Analysis (ATHEANA) - Technical Basis and Methodology Description,* DRAFT NUREG/CR-6350, to be published.

- /4/ Cooper, S.E. et al, *Knowledge-Base for the New Human Reliability Analysis Method, "A Technique for Human Error Analysis" (ATHEANA)*, in the proceedings of PSA 96, Park City, Utah, September 1996.
- /5/ Parry, G.W., Luckas, W.J., Wreathall, J., Cooper, S.E., and Bley, D.C., *Process Description for ATHEANA: A Technique for Human Error Analysis*, Brookhaven National Laboratory Technical Report, L-2415195-2, December 30, 1995.
- /6/ Roth, E.M., Mumaw, R.J., and Lewis, P.M., *An Empirical Investigation of Operator Performance in Cognitively Demanding Simulated Emergencies*, NUREG/CR-6208, Westinghouse Science and Technology Center: Pittsburgh, PA, July 1994.
- /7/ Julius, J.A., Jorgenson, E.M., Parry, G.W., and Mosleh, A.M., "A Procedure for the Analysis of Errors of Commission in a Probabilistic Safety Assessment of a Nuclear Power Plant at Full Power" *Reliability Engineering and System Safety*, Vol. 50, (1995), pages 189-201.

Appendix F is contained in a separate volume.